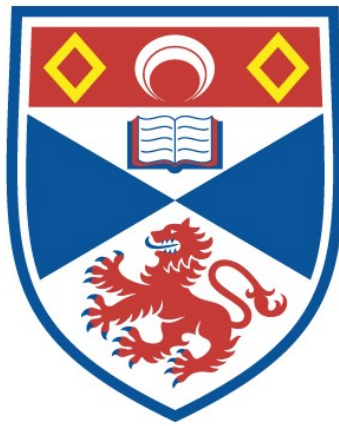


**TRANSITING EXTRA-SOLAR PLANETS IN THE FIELD OF  
OPEN CLUSTER NGC 7789**

**Daniel Martyn Bramich**

**A Thesis Submitted for the Degree of PhD  
at the  
University of St Andrews**



**2005**

**Full metadata for this item is available in  
St Andrews Research Repository  
at:**

**<http://research-repository.st-andrews.ac.uk/>**

**Please use this identifier to cite or link to this item:**

**<http://hdl.handle.net/10023/12944>**

**This item is protected by original copyright**

85  
L

Transiting Extra-Solar Planets In The Field Of Open Cluster  
NGC 7789

Submitted for the degree of Ph.D. by Daniel Martyn Bramich

January 13, 2005



ProQuest Number: 10171060

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10171060

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

TH E768



## DECLARATION

I, Daniel Martyn Bramich, hereby certify that this thesis, which is approximately 40000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date 13/01/05 Signature Of Candidate ..

I was admitted as a research student in September 2001 and as a candidate for the degree of Ph.D. in September 2001; the higher study for which this is a record was carried out in the University of St. Andrews between 2001 and 2004.

Date 13/01/05 Signature Of Candidate .

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 13 Jan 05 Signature Of Supervisor ..

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

Date 13/01/05 Signature Of Candidate .

---

## ABSTRACT

We present results from 30 nights of observations of the intermediate-age Solar-metallicity open cluster NGC 7789 with the WFC camera on the INT telescope in La Palma. From  $\sim 900$  epochs, we obtained lightcurves and Sloan  $r' - i'$  colours for  $\sim 33000$  stars, with  $\sim 2400$  stars with better than 1% precision. We find 24 transit candidates, 14 of which we can assign a period. We rule out the transiting planet model for 21 of these candidates using various robust arguments. For 2 candidates we are unable to decide on their nature, although it seems most likely that they are eclipsing binaries as well. We have one candidate exhibiting a single eclipse for which we derive a radius of  $1.81^{+0.09}_{-0.00} R_J$ . Three candidates remain that require follow-up observations in order to determine their nature.

Monte Carlo simulations reveal that we expected to detect  $\sim 2$  transiting 3d to 5d hot Jupiter planets from all the stars in our sample if 1% of stars host such a companion and that a typical hot Jupiter radius is similar to that of HD 209458b. Our failure to find good transiting hot Jupiter candidates allows us to place an upper limit on the 3d to 5d hot Jupiter fraction of 2.6% for all stars at the 1% significance level, and similar useful limits on the hot Jupiter fraction of the different star types in our sample.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	The Discovery Of Extra-Solar Planets . . . . .	1
1.2	Characteristics Of The Extra-Solar Planets . . . . .	4
1.2.1	The Mass And Period Distributions . . . . .	11
1.2.2	The Orbital Eccentricity And Period Correlation . . . . .	14
1.2.3	Hot Jupiters And Their Properties . . . . .	16
1.2.4	Planet Frequency And Stellar Metallicity . . . . .	19
1.2.5	The Variety Of Extra-Solar Planets . . . . .	22
1.3	Star And Planetary Formation . . . . .	22
1.3.1	Star Formation . . . . .	22
1.3.2	Planet Formation . . . . .	24
1.3.3	Planet Migration . . . . .	26
1.4	Summary . . . . .	27
<b>2</b>	<b>A Menagerie Of Detection Techniques</b>	<b>28</b>
2.1	Radial Velocity . . . . .	29
2.2	Direct Imaging And Reflected Light . . . . .	30
2.3	Gravitational Microlensing . . . . .	32
2.4	Transit Photometry . . . . .	35

---

2.4.1	Transit Probability . . . . .	35
2.4.2	Transit Duration . . . . .	38
2.4.3	Transit Lightcurve Morphology . . . . .	39
2.4.4	Transit Signal To Noise . . . . .	44
2.4.5	The History Of The Transit Technique . . . . .	47
2.4.6	Transit Mimics And Follow-Up Strategies . . . . .	49
2.4.7	Conclusions . . . . .	52
2.5	Summary . . . . .	53
<b>3</b>	<b>A Transit Survey Of The Field Of Open Cluster NGC 7789</b>	<b>55</b>
3.1	Introduction . . . . .	56
3.2	Observations . . . . .	56
3.3	CCD Reductions . . . . .	58
3.3.1	Preprocessing: CCD Calibrations . . . . .	58
3.3.2	Photometry: Difference Image Analysis . . . . .	62
3.4	Astrometry And Colour Data . . . . .	66
3.4.1	Astrometry . . . . .	66
3.4.2	Colour Indices . . . . .	67
3.4.3	Colour Magnitude Diagrams . . . . .	67
3.4.4	Stellar Masses, Radii And Distances . . . . .	71
3.4.5	Number Of Expected Transiting Planets . . . . .	75
3.5	Summary . . . . .	78
<b>4</b>	<b>21 Eclipsing Binaries And 3 Planetary Transit Candidates</b>	<b>79</b>
4.1	Transit Detection . . . . .	79
4.2	Transit Candidates . . . . .	83
4.2.1	Theoretical Models . . . . .	83

4.2.2	Lightcurve Modelling Procedure . . . . .	84
4.2.3	Eclipsing Binaries With Undetermined Periods . . . . .	92
4.2.4	Eclipsing Binaries Exhibiting Secondary Eclipses . . . . .	97
4.2.5	Eclipsing Binaries Exhibiting Ellipsoidal Variations And Heating Effects . . . . .	100
4.2.6	A Possible Long Period Cataclysmic Variable . . . . .	101
4.2.7	More Eclipsing Binaries . . . . .	102
4.2.8	INT-7789-TR-1 (Star 45134) . . . . .	109
4.2.9	INT-7789-TR-2 (Star 46691) . . . . .	109
4.2.10	INT-7789-TR-3 (Star 49512) . . . . .	114
4.2.11	Finding Charts . . . . .	116
4.3	Summary . . . . .	118
<b>5</b>	<b>Limits On The Hot Jupiter Fraction In The Field Of NGC 7789</b>	<b>119</b>
5.1	Introduction . . . . .	119
5.2	Detection Probabilities And False Alarm Rates . . . . .	120
5.3	Monte Carlo Simulations . . . . .	122
5.4	Number Of Expected Transiting Planets . . . . .	127
5.5	Number Of Expected Transiting Hot Jupiters . . . . .	131
5.6	Discussion . . . . .	138
5.7	Summary . . . . .	141
<b>6</b>	<b>Conclusions</b>	<b>142</b>
<b>A</b>	<b>The <math>M \sin i</math> Ambiguity</b>	<b>145</b>
<b>B</b>	<b>Useful Data/Constants</b>	<b>150</b>
<b>C</b>	<b>Radial Velocity Curves</b>	<b>152</b>

# List of Tables

1.1	Extra-solar planet catalogue . . . . .	6
1.2	Results of the fits to the ECDFs . . . . .	18
1.3	Core accretion timescales and particle/object sizes . . . . .	25
3.1	Properties of the open cluster NGC 7789 . . . . .	57
3.2	Readout noise and gain for each chip . . . . .	59
3.3	Magnitude offsets between runs . . . . .	73
3.4	Number of stars with lightcurves and colours for each chip . . . . .	73
3.5	Magnitude offsets used to calibrate the instrumental magnitudes . . . . .	73
4.1	Number of raw transit candidates and remaining transit candidate lightcurves	81
4.2	Star and lightcurve properties for the transit candidates . . . . .	86
4.3	Lightcurve and companion properties for the transit candidates . . . . .	89
4.4	Lightcurve and companion properties for the transit candidates (continued)	90
4.5	Star, companion and lightcurve properties for INT-7789-TR-3 (star 49512) .	113
5.1	The different subsets of stars used when calculating the number of expected transiting planets and false alarms . . . . .	127
5.2	Results of the Monte Carlo simulations . . . . .	136

# List of Figures

1.1	Mean mass histogram . . . . .	11
1.2	Period histogram and cumulative distribution function . . . . .	12
1.3	Mean mass and period correlation . . . . .	14
1.4	Orbital eccentricity and period correlation . . . . .	15
1.5	Empirical cumulative distribution functions . . . . .	17
1.6	Planet frequency against stellar metallicity . . . . .	21
2.1	General configuration of a gravitational lens . . . . .	32
2.2	Configuration of an extra-solar planetary system at the time of a touching (just grazing) eclipse . . . . .	36
2.3	General configuration of an extra-solar planetary system at the start of the transit ingress . . . . .	36
2.4	Transit lightcurve morphology as a function of planet type . . . . .	41
2.5	Transit lightcurve morphology as a function of inclination and the linear limb darkening coefficient . . . . .	43
2.6	Typical photometric observations of a transit event . . . . .	46
3.1	Variance versus signal for each chip . . . . .	60
3.2	Diagnostic data for Chip 4 . . . . .	61
3.3	RMS diagrams for each chip . . . . .	65

---

3.4	Instrumental colour magnitude diagrams for each chip . . . . .	68
3.5	Instrumental CMD for Chip 4 . . . . .	69
3.6	Plots of Sloan $r'$ magnitude and lightcurve RMS against stellar radius . . .	72
3.7	Mass-radius relationship for the theoretical main sequence . . . . .	75
3.8	Theoretical and observed main sequence relations in the $M_R$ versus $R - I$ domain . . . . .	76
4.1	An example boxcar transit fit . . . . .	80
4.2	Plot of the transit statistic against the out-of-transit reduced chi squared .	82
4.3	Eclipsing binaries with undetermined periods . . . . .	93
4.4	Eclipsing binaries with undetermined periods . . . . .	94
4.5	Eclipsing binaries with undetermined periods . . . . .	95
4.6	Eclipsing binaries exhibiting secondary eclipses . . . . .	98
4.7	Eclipsing binaries exhibiting ellipsoidal variations and heating effects . . . .	99
4.8	Possible long period cataclysmic variable . . . . .	102
4.9	Stars 6995 and 7628 . . . . .	104
4.10	Stars 22738 and 50313 . . . . .	105
4.11	Stars 64804 and 73852 . . . . .	106
4.12	Planetary transit candidates 45134, 46691 and 49512 . . . . .	110
4.13	Eclipsing binary fits for star 49512 . . . . .	111
4.14	Finding charts . . . . .	117
5.1	Detection probability and false alarm rate as functions of detection threshold and period for a single star . . . . .	125
5.2	Number of expected transiting planets and false alarms for all stars as func- tions of the detection threshold for two different periods . . . . .	129
5.3	Number of expected transiting planets and false alarms for all stars as func- tions of the period . . . . .	130



---

5.4	Number of expected transiting planets and false alarms for all stars as functions of the detection threshold for different planet types . . . . .	133
5.5	Number of expected transiting planets as a function of the detection threshold for different star types and planet types . . . . .	134
5.6	The upper limit on the planet fraction as a function of star type and planet type . . . . .	139
A.1	General configuration of an extra-solar planet in a circular orbit . . . . .	146
A.2	Definition of the $xy$ -plane for the extra-solar planet . . . . .	146
C.1	General configuration of the orbit of a star about its centre of mass . . . . .	153

# 1

---

## Introduction

### 1.1 The Discovery Of Extra-Solar Planets

Some of the first steps towards answering one of humanity's most fundamental questions about the Universe, "Are we alone?", lie in the search for planets that orbit stars other than the Sun. Such planets are termed extra-solar planets (or exoplanets). Earth is the only life-bearing planet that we know of and we might expect other solar systems to be similar to our own by assuming that we are not special in any way. Once we know of other planetary systems we may start to look for evidence of life through the detection of biomarkers, chemical elements associated with life processes on Earth, in the spectra of light from an exoplanet. Any speculation on what this life may actually be like lies in the realm of science fiction for the near future.

Until recently the subject of extra-solar planets also lay in the realm of fiction rather than fact. Hard evidence that these objects existed was only presented for the first time in 1992 with a rather unexpected detection. Wolszczan & Frail (1992) identified periodic variations ( $\pm 2\text{ms}$ ) in the time of arrival of pulses received from the 6.2ms pulsar PSR B1257+12. This led to the conclusion that two planets of masses similar to that of the Earth were orbiting the pulsar at distances similar to that of Mercury from the Sun. Further work revealed another even lower mass planet orbiting the pulsar and application of non-Keplerian

dynamics allowed the measurement of the true masses and orbital inclinations of the inner two planets (Wolszczan 1994; Konacki, Maciejewski, & Wolszczan 2000; Konacki & Wolszczan 2003). Pulsars (rapidly rotating neutron stars) are formed during supernovae. The detected planets may have survived a supernova explosion or they may have formed from an accretion disk after the supernova phase. However, both of these evolutionary paths and the hostile pulsar environment imply that these planets are unlikely to harbour life (as we know it!).

The defining breakthrough came in November 1995 with the first detection by Mayor & Queloz (1995) of a Jupiter mass planet orbiting a main sequence star, 51 Pegasi. By measuring the radial velocity of the host star during four different epochs, periodic variations were detected that could be fitted adequately by the presence of a  $0.47M_J/\sin i$  mass planet orbiting in a 4.23 day circular orbit of radius 0.05AU, where  $M_J$  is the mass of Jupiter and  $i$  is the orbital inclination to the line of sight ( $i = 90^\circ$  when the line of sight lies in the plane of the orbit). It is impossible to determine the value of  $\sin i$  by use of the radial velocity technique alone, and therefore the mass derived for the planetary companion serves as a lower limit. Having said this, assuming that no particular orbital inclination is preferred, then with greater than 99% probability the mass of the companion is  $\leq 3.5M_J$  (see Appendix A, Example 1). By considering constraints originating from the observed rotational velocity of 51 Pegasi (Soderblom 1983; Baranne, Mayor, & Poncet 1979) and its low chromospheric emission (Noyes et al. 1984), Mayor & Queloz (1995) were able to give an upper limit of  $2M_J$  for the mass of the companion.

This discovery was unexpected due to the high mass of the planetary companion orbiting the star at such a small distance and it sparked a controversy about the nature of the radial velocity variations. Some authors claimed to have detected 4.23 day periodic variations in the shapes of various absorption lines from high resolution spectroscopy of the star 51 Pegasi (Hatzes, Cochran, & Johns-Krull 1997; Gray & Hatzes 1997) which could not be explained by the orbiting planet hypothesis. Further, it was claimed that low order and low degree nonradial oscillations could fully account for the radial velocity observations and the changes in the shape of the line profiles. Brown et al. (1998) refuted these claims with their own spectroscopic observations which did not show any variations in the line profiles.

Meanwhile Marcy et al. (1997) published more radial velocity measurements showing the 4.23 day periodicity and stating that the only viable interpretation was one of a Jupiter mass companion. However, a general consensus was reached that the radial velocity variations were most consistent with a planetary companion with the publication by Hatzes, Cochran, & Bakker (1998) of a lack of spectral variability in 51 Pegasi, refuting their own previous claims.

Soon after the discovery of 51 Pegasi b (the letter b denotes the reference to the planet rather than the star), other groups announced more planet candidates from radial velocity (RV) surveys (Butler & Marcy 1996; Marcy & Butler 1996; Butler et al. 1997) leading to an explosion in the number of planetary detections. With an ever increasing time baseline for the RV measurements, it has been possible to detect planets of longer periods, consequently probing larger orbital distances from the host stars. The  $M_p \sin i$  detection limit is determined by the precision of the RV measurements which typically ranges from  $\sim 10 \text{ms}^{-1}$  for the smaller telescopes (Baranne et al. 1996) to  $\sim 3 \text{ms}^{-1}$  for the larger telescopes (Tinney et al. 2001). Improvements in the efficiency of the spectrographs employed in the detection of extra-solar planets has increased the precision obtained from the RV measurements to around  $\sim 1\text{-}2 \text{ms}^{-1}$  (Pepe et al. 2004). However, there seems to be a fundamental limit to the attainable precision defined by the intrinsic velocity stability of the target stars (Saar, Butler, & Marcy 1998; Saar & Fischer 2000). Such RV variations, commonly called “jitter”, are induced by the rotation of star spots and/or convective inhomogeneities and their temporal evolution. It should also be noted that the RV technique is limited to surveying bright target stars in order to provide the necessary high signal to noise (S/N) spectra. As a consequence, this limits the RV surveys to Solar neighbourhood stars. There has also been a tendency to target main sequence stars similar to our Sun in the quest to find a Solar System analogue. For instance, the Anglo-Australian planet search targets F, G and K main sequence stars down to  $V=7.5$  mag (Jones 2002).

By March 2000, 34 extra-solar planets were known from RV measurements and recently detections at a rate of  $\sim 20$  planets per year have been the norm, almost exclusively discovered by the RV method. The confirmed 123 planets to date (01/08/2004) are listed in an extra-solar planet encyclopedia at the web address:

<http://www.obspm.fr/encycl/cat1.html>

## 1.2 Characteristics Of The Extra-Solar Planets

At this stage, clarification is required as to exactly what class of objects are defined to be planets. An object with a mass greater than  $\sim 0.08 M_{\odot} \equiv 84 M_J$  has a core temperature high enough to ignite the thermonuclear fusion of hydrogen, and hence to self luminesce. This object is a star and stars are thought to form from the collapse of rotating interstellar gas and dust clouds via gravitational instability (Boss 1980). The minimum mass required to form an object via the gravitational collapse of such a gas cloud is thought to lie in the range  $7-20 M_J$  (Boss 1986). Objects formed this way with a mass less than  $\sim 0.08 M_{\odot}$  are referred to as brown dwarfs (Tarter 1986; Burrows & Liebert 1993). They emit radiation mainly in the infrared owing to the thermal energy of their creation and for brown dwarfs with masses greater than  $\sim 12 M_J$ , deuterium fusion in their cores will contribute to their luminescence.

Planets are thought to form via the agglomeration and accretion of material within the gas and dust disk of a protostar, and planets do not luminesce via thermonuclear fusion at any stage during their lifetimes. The lower mass limit of  $\sim 12 M_J$  for deuterium fusion depends weakly on various factors (chemical composition etc.) and therefore a safe upper mass limit for planets of  $\sim 10 M_J$  will be adopted here. It must be pointed out that since brown dwarfs may also form as stellar companions, there lies a “grey” area between what constitutes a planet and what constitutes a brown dwarf.

Table 1.1 lists the main properties of the extra-solar planets orbiting main sequence stars discovered to date. These data are taken from the extra-solar planet encyclopedia mentioned in Chapter 1.1, excluding 4 objects with inaccurately determined values of  $M_p \sin i$  and excluding 6 objects with  $M_p \sin i \geq 10 M_J$ . We also exclude the planet OGLE-235/MOA-53b which has incomplete data. The resulting catalogue presented in Table 1.1 consists of 112 confirmed planets in 103 planetary systems, a statistically significant sample. It is this sample that we use in the analysis of the following sections where we describe the distributions of the various properties of the extra-solar planets which have started to reveal

a fine structure that until very recently was not apparent.

Table 1.1: Extra-solar planet catalogue ordered by increasing period of innermost planetary companion.  
 Col. 4: Planetary semi-major axis (AU). Col. 5: Orbital period (days). Col. 6: Orbital eccentricity.

Planet No.	Star Name	$M_p \sin i (M_J)$	$a_p$ (AU)	$P$ (d)	$e$
1	OGLE-TR-56	1.45*	0.0225	1.2	0.0
2	OGLE-TR-113	1.35*	0.0228	1.43	0.0
3	OGLE-TR-132	1.01*	0.0306	1.69	0.0
4	HD 73256	1.85	0.037	2.54863	0.038
5	HD 83443	0.41	0.04	2.985	0.08
6	HD 46375	0.249	0.041	3.024	0.04
7	HD 179949	0.84	0.045	3.093	0.05
8	HD 187123	0.52	0.042	3.097	0.03
9	$\tau$ Boo	3.87	0.0462	3.3128	0.018
10	HD 330075	0.76	0.043	3.369	0.0
11	BD-10 3166	0.48	0.046	3.487	0.0
12	HD 75289	0.42	0.046	3.51	0.054
13	HD 209458	0.69*	0.045	3.524738	0.0
14	HD 76700	0.197	0.049	3.971	0.0
15	51 Peg	0.468	0.052	4.23077	0.0
16	$\nu$ And	0.69	0.059	4.6170	0.012
17		1.19	0.829	241.5	0.28
18		3.75	2.53	1284	0.27
19	HD 49674	0.12	0.0568	4.948	0.0
20	HD 68988	1.90	0.071	6.276	0.14
21	HD 168746	0.23	0.065	6.403	0.081
22	HD 217107	1.28	0.07	7.11	0.14
23	HD 130322	1.08	0.088	10.724	0.048
continued on next page					

<i>continued from previous page</i>					
Planet No.	Star Name	$M_p \sin i (M_J)$	$a_p$ (AU)	$P$ (d)	$e$
24	HD 108147	0.41	0.104	10.901	0.498
25	HD 38529	0.78	0.129	14.309	0.29
26	55 Cnc	0.84	0.11	14.65	0.02
27		4.05	5.9	5360	0.16
28	Gl 86	4.0	0.11	15.78	0.046
29	HD 195019	3.43	0.14	18.3	0.05
30	HD 6434	0.48	0.15	22.09	0.30
31	HD 192263	0.72	0.15	24.348	0.0
32	Gliese 876	0.56	0.13	30.1	0.12
33		1.98	0.21	61.02	0.27
34	$\rho$ CrB	1.04	0.22	39.845	0.04
35	HD 74156	1.86	0.294	51.643	0.636
36	HD 37605	2.85	0.26	55.2	0.736
37	HD 168443	7.7	0.29	58.116	0.529
38	HD 3651	0.2	0.284	62.23	0.63
39	HD 121504	0.89	0.32	64.6	0.13
40	HD 178911 B	6.292	0.32	71.487	0.1243
41	HD 16141	0.23	0.35	75.560	0.28
42	HD 80606	3.41	0.439	111.78	0.927
43	70 Vir	7.44	0.48	116.689	0.4
44	HD 216770	0.65	0.46	118.45	0.37
45	HD 52265	1.13	0.49	118.96	0.29
46	GJ 3021	3.21	0.49	133.82	0.505
47	HD 37124	0.75	0.54	152.4	0.10
48		1.2	2.5	1495	0.69
49	HD 219449	2.9	0.3	182	0.0
<i>continued on next page</i>					



continued from previous page					
Planet No.	Star Name	$M_p \sin i$ ( $M_J$ )	$a_p$ (AU)	$P$ (d)	$e$
50	HD 73526	3.0	0.66	190.5	0.34
51	HD 104985	6.3	0.78	198.2	0.03
52	HD 82943	0.88	0.73	221.6	0.54
53		1.63	1.16	444.6	0.41
54	HD 169830	2.88	0.81	225.62	0.31
55		4.04	3.60	2102	0.33
56	HD 8574	2.23	0.76	228.8	0.40
57	HD 89744	7.99	0.89	256.6	0.67
58	HD 134987	1.58	0.78	260	0.25
59	HD 12661	2.30	0.83	263.6	0.096
60		1.57	2.56	1444.5	0.1
61	HD 150706	1.0	0.82	264.9	0.38
62	HD 40979	3.32	0.811	267.2	0.25
63	HD 59686	6.5	0.8	303	0.0
64	HD 810	2.26	0.925	320.1	0.161
65	HD 142	1.36	0.980	338.0	0.37
66	HD 92788	3.8	0.94	340	0.36
67	HD 28185	5.6	1.0	385	0.06
68	HD 142415	1.62	1.05	386.3	0.5
69	HD 177830	1.28	1.00	391	0.43
70	HD 4203	1.65	1.09	400.944	0.46
71	HD 108874	1.65	1.07	401	0.20
72	HD 128311	2.63	1.06	414	0.21
73	HD 27442	1.28	1.18	423.841	0.07
74	HD 210277	1.28	1.097	437	0.45
75	HD 19994	2.0	1.3	454	0.2
continued on next page					

<i>continued from previous page</i>					
Planet No.	Star Name	$M_p \sin i (M_J)$	$a_p$ (AU)	$P$ (d)	$e$
76	HD 20367	1.07	1.25	500	0.23
77	HD 114783	0.9	1.20	501.0	0.1
78	HD 147513	1.0	1.26	540.4	0.52
79	HIP 75458	8.64	1.34	550.651	0.71
80	HD 222582	5.11	1.35	572.0	0.76
81	HD 65216	1.21	1.37	613.1	0.41
82	HD 160691	1.7	1.5	638	0.31
83	HD 141937	9.7	1.52	653.22	0.41
84	HD 41004A	2.3	1.31	655	0.39
85	HD 47536	4.96	1.61	712.13	0.20
86	HD 23079	2.61	1.65	738.459	0.10
87	16 Cyg B	1.69	1.67	798.938	0.67
88	HD 4208	0.80	1.67	812.197	0.05
89	HD 114386	0.99	1.62	872	0.28
90	$\gamma$ Cephei	1.59	2.03	902.96	0.2
91	HD 213240	4.5	2.03	951	0.45
92	HD 10647	0.91	2.10	1040	0.18
93	HD 10697	6.12	2.13	1077.906	0.11
94	47 Uma	2.41	2.10	1095	0.096
95		0.76	3.73	2594	0.1
96	HD 190228	4.99	2.31	1127	0.43
97	HD 114729	0.82	2.08	1131.478	0.31
98	HD 111232	6.8	1.97	1143	0.20
99	HD 2039	4.85	2.19	1192.582	0.68
100	HD 50554	4.9	2.38	1279.0	0.42
101	HD 196050	3.0	2.5	1289	0.28
<i>continued on next page</i>					

<i>continued from previous page</i>					
Planet No.	Star Name	$M_p \sin i$ ( $M_J$ )	$a_p$ (AU)	$P$ (d)	$e$
102	HD 216437	2.1	2.7	1294	0.34
103	HD 216435	1.49	2.7	1442.919	0.34
104	HD 106252	6.81	2.61	1500	0.54
105	HD 23596	7.19	2.72	1558	0.314
106	14 Her	4.74	2.80	1796.4	0.338
107	HD 72659	2.55	3.24	2185	0.18
108	HD 70642	2.0	3.3	2231	0.1
109	HD 33636	9.28	3.56	2447.292	0.53
110	$\epsilon$ Eridani	0.86	3.3	2502.1	0.608
111	HD 30177	9.17	3.86	2819.654	0.30
112	Gl 777A	1.33	4.8	2902	0.48

\*The mass  $M_p$  of the extra-solar planet is listed instead of the value of  $M_p \sin i$ .

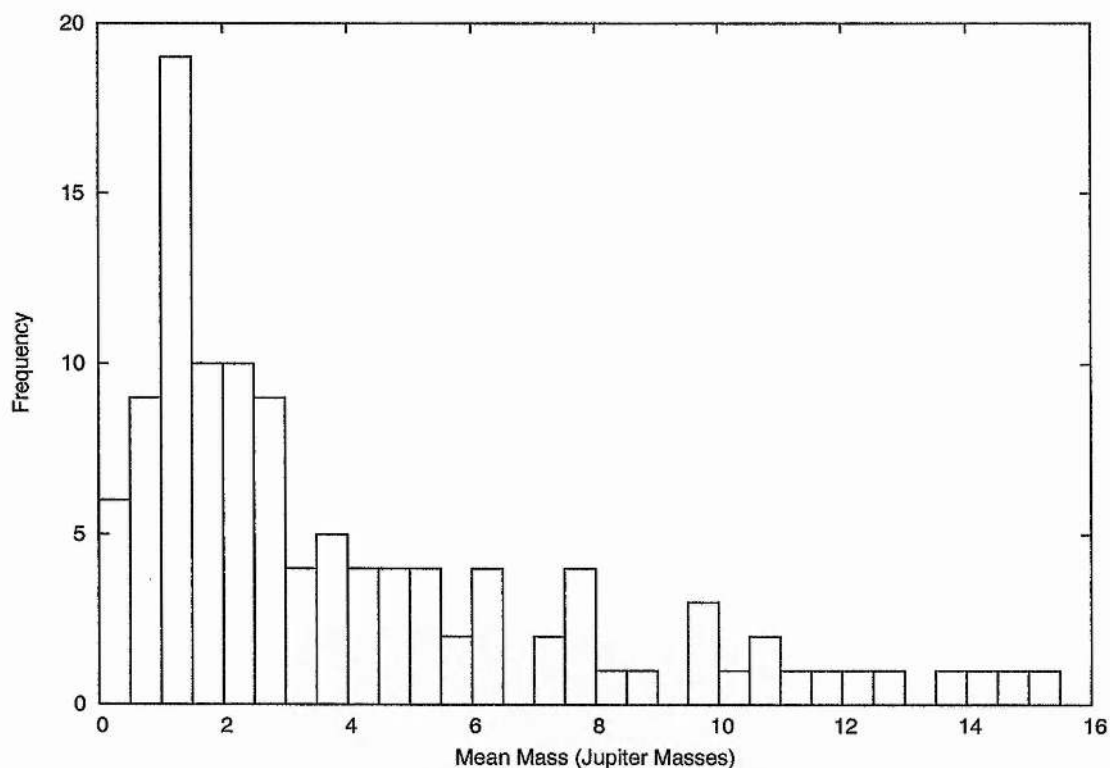
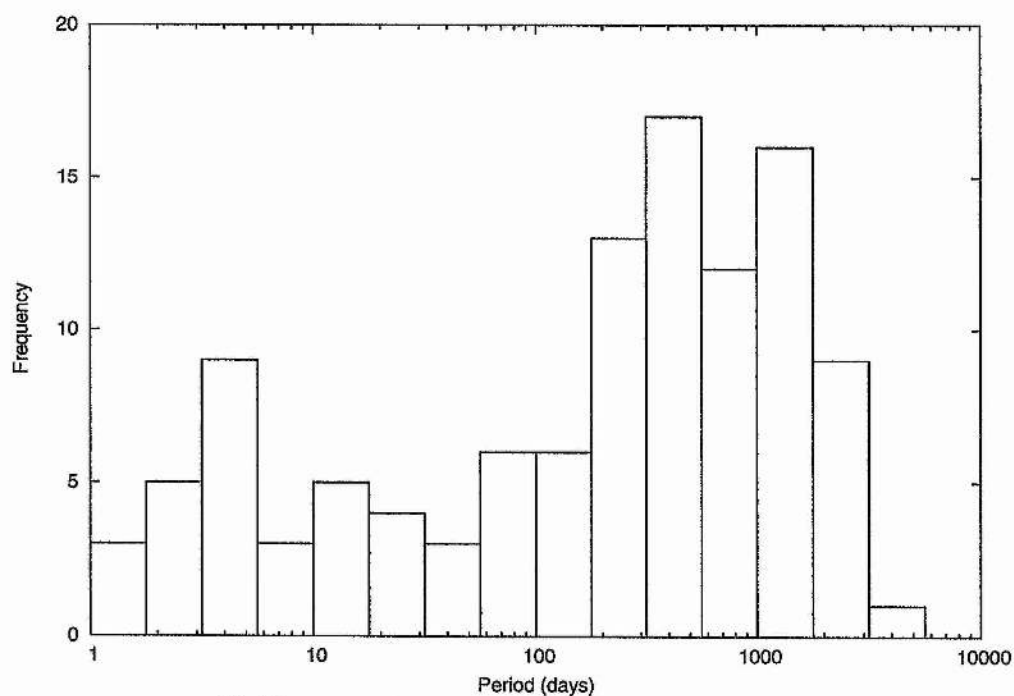


Figure 1.1: Histogram of the mean mass of the extra-solar planets.

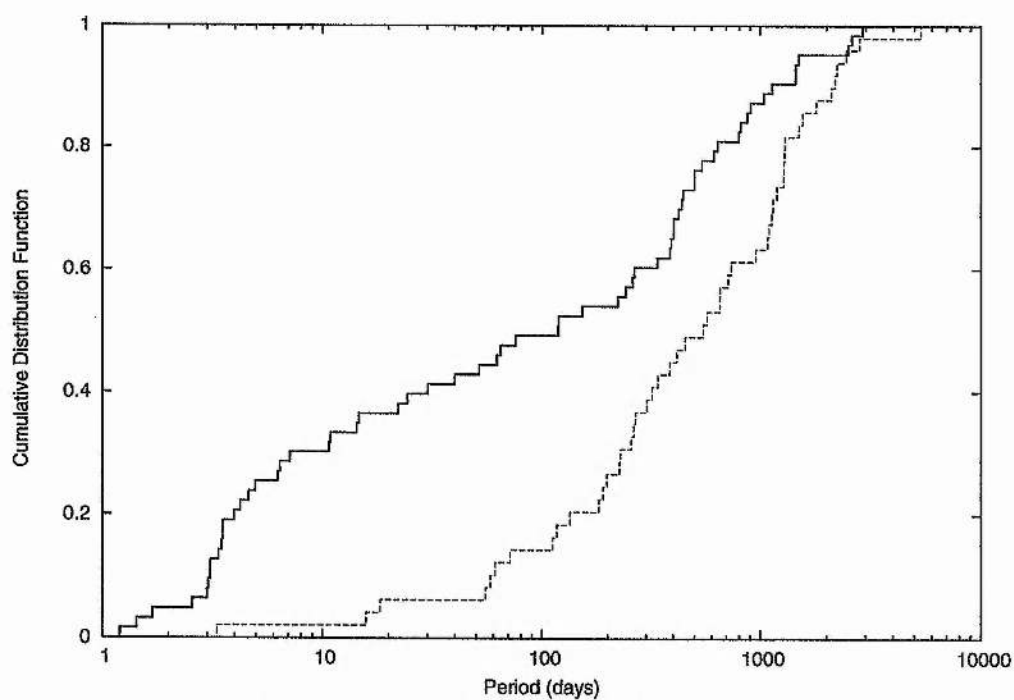
### 1.2.1 The Mass And Period Distributions

Radial velocity measurements of a planet host star supply the observable quantity  $M_p \sin i$  as a lower limit to the mass of the extra-solar planet (see Section 2.1). Making the assumption that the orbital orientation of the extra-solar planet is random means that we can calculate the mean mass  $\langle M_p \rangle$  of the extra-solar planet as  $\langle M_p \rangle = \frac{\pi}{2} M_p \sin i$  (see Appendix A, Example 2). Figure 1.1 shows a histogram of the mean mass distribution for the extra-solar planets listed in Table 1.1. The mean mass distribution rises towards lower masses down to  $\sim 1M_J$ . The failure to rise any further for even lower masses is due the limited detectability of these planets using the RV technique.

Figure 1.2(a) shows a histogram of the period distribution for the extra-solar planets.



(a) Histogram of the period of the extra-solar planets.



(b) Cumulative distribution function against period for the less massive extra-solar planets ( $\langle M_p \rangle \leq 3M_J$  with the continuous line) and for the more massive extra-solar planets ( $\langle M_p \rangle > 3M_J$  with the dashed line).

Figure 1.2: Period histogram and cumulative distribution function

The period distribution shows a sharp cut off at around  $\sim 3$  days, especially if we ignore the 3 recently discovered OGLE planets that seem to form part of a new class of planets (“very hot Jupiters”). The period distribution then drops towards higher periods and rises abruptly again for periods  $\gtrsim 100$  days, revealing a “period valley” for periods between  $\sim 10$  and  $\sim 100$  days (Udry, Mayor, & Santos 2003). Observationally, this may be explained by the combination of two different planet populations, a lower mass planet population that peaks at short periods and extends to longer periods, and a higher mass planet population that almost exclusively has periods  $\gtrsim 100$  days. The existence of the two populations is illustrated clearly in Figure 1.2(b) where we plot the cumulative distribution function (CDF) against period for two planet populations, the less massive extra-solar planets ( $\langle M_p \rangle \leq 3M_J$  with the continuous line) and the more massive extra-solar planets ( $\langle M_p \rangle > 3M_J$  with the dashed line). The statistical significance of the lack of higher mass planets on short periods has been verified by Zucker & Mazeh (2002) and Udry, Mayor, & Santos (2003) using the fact that if such planets existed then they should be easily detectable by the RV technique.

Figure 1.3 shows a scatter plot of the mean mass against period for the extra-solar planets. The rectangular region in the upper left of the diagram delimited by the dashed line highlights the region with few higher mass planets at short periods ( $\langle M_p \rangle > 3M_J$  and  $P \leq 100$  days). In fact, if you remove the planets that orbit in multiple-star systems (circles with dots), then this region becomes almost devoid of data points, suggesting that planetary formation and evolution in multiple-star systems could follow different paths to those in single-star systems. The other rectangular region in the lower right of the diagram, also delimited by a dashed line, reveals a lack of lower mass planets at long periods ( $\langle M_p \rangle \leq 1M_J$  and  $P \geq 100$  days). This feature could be related to the observational bias from the RV technique that lower mass and longer period planets are more difficult to detect. However, Monte-Carlo simulations by Udry, Mayor, & Santos (2003) show that this region is indeed devoid of planets with a 99.97% confidence level.

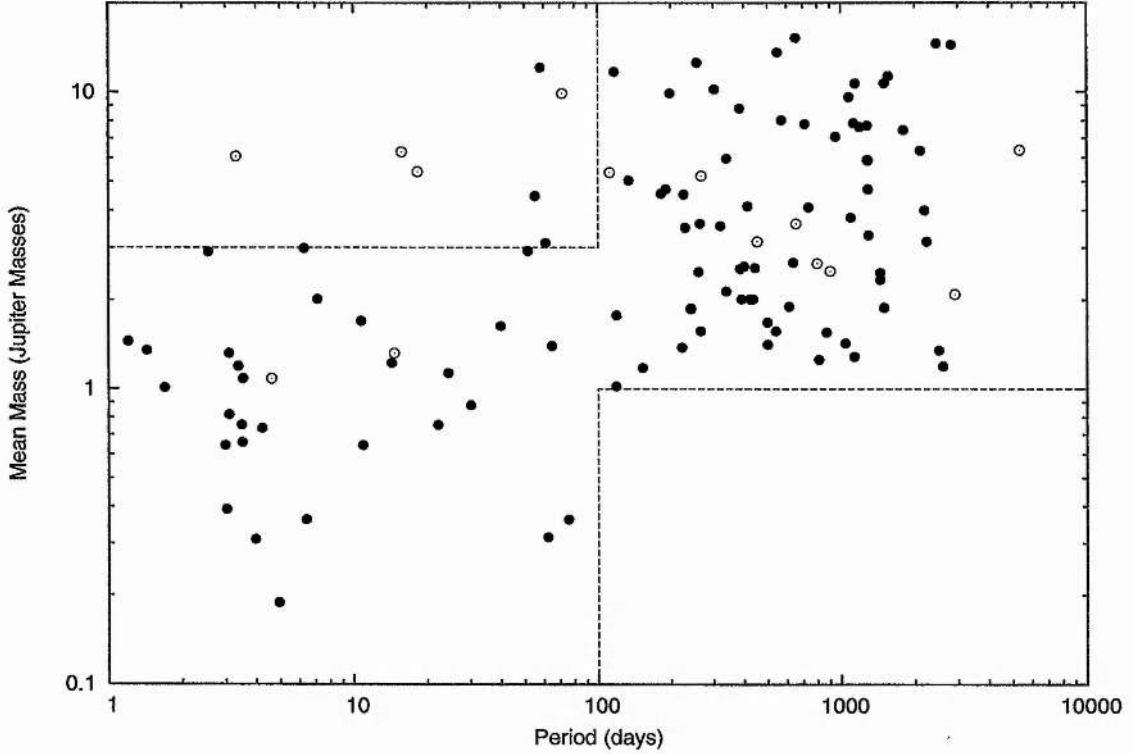


Figure 1.3: Plot of mean mass against period for the extra-solar planets. Filled circles represent planets that orbit in single-star systems. Circles with dots represent planets that orbit in multiple-star systems.

### 1.2.2 The Orbital Eccentricity And Period Correlation

The extra-solar planets with the smallest orbital semi-major axes (shortest periods) are likely to have undergone orbital circularisation via tidal interaction with the host star. A scatter plot of orbital eccentricity against period (Figure 1.4) clearly shows that all extra-solar planets with periods of less than 6.0 days have orbital eccentricities of less than 0.1 (lower left rectangular region delimited by a dashed line) indicative of circular orbits. Removing these planets from our sample and calculating the Spearman's rank correlation coefficient between orbital eccentricity  $e$  and period  $P$  yields  $r_{95} = 0.2277 \dots \approx 0.228$ . Assuming a null-hypothesis that  $e$  and  $P$  are uncorrelated allows one to calculate, for a given  $N$ , the probability of obtaining a value of  $|r_N|$  greater than or equal to  $k$  where  $k \in \mathbb{R}$ . This probability, denoted by  $P(|r_N| \geq k)$ , is called the statistical significance of

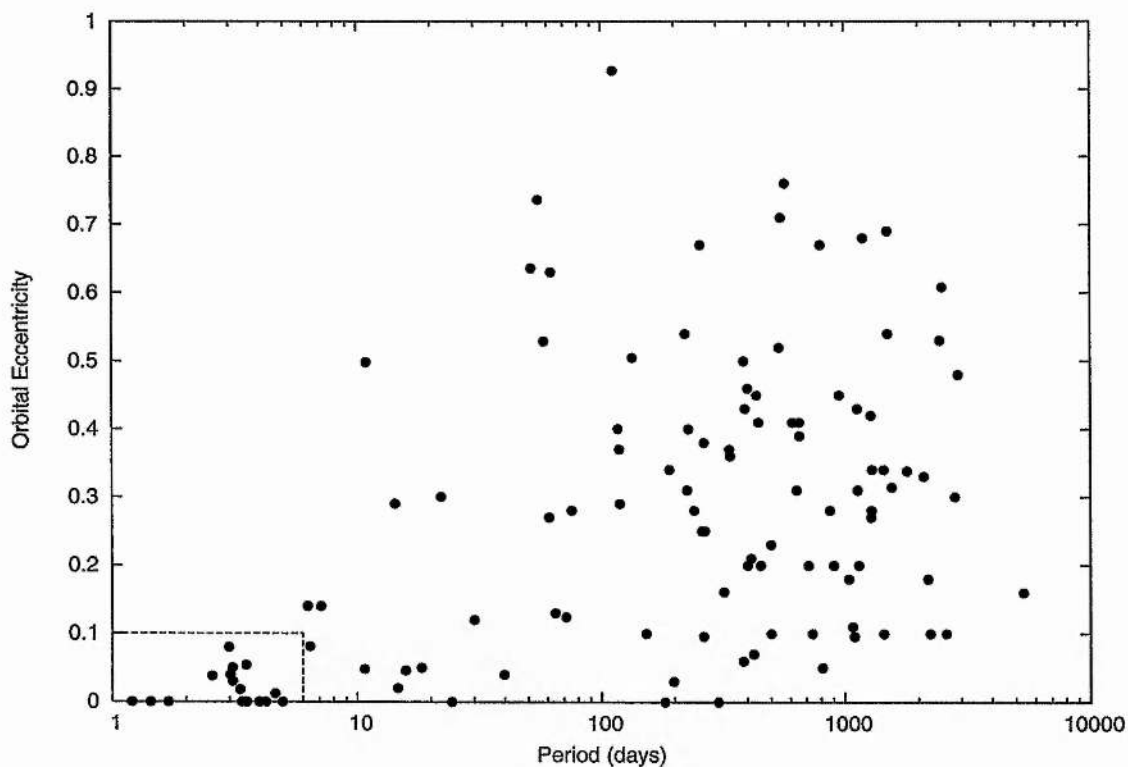


Figure 1.4: Plot of orbital eccentricity against period for the extra-solar planets.

the test and the smaller the value of  $P(|r_N| \geq k)$ , the less likely we are to reject the null-hypothesis when it is actually correct (Type I error). If we choose a significance level  $\alpha$  and we have  $P(|r_N| \geq k) \leq \alpha$ , then we can reject the null-hypothesis at the  $\alpha$  significance level. Since  $P(|r_N| \geq 0.228) = 0.0264 \dots \leq 0.05$  for  $N = 95$ , we can reject the null-hypothesis that  $e$  and  $P$  are uncorrelated at the  $\alpha = 0.05$  significance level, and reasonably conclude that  $e$  and  $P$  are indeed correlated in some way. As an aside, it is interesting to note that without removing the extra-solar planets with  $P \leq 6.0$  days from the sample,  $r_{112} = 0.4888 \dots \approx 0.489$  and  $P(|r_N| \geq 0.489) = 4.54 \dots \times 10^{-8} \leq 1.0 \times 10^{-7}$  for  $N = 112$ .



### 1.2.3 Hot Jupiters And Their Properties

The discovery of close-in Jupiter-mass companions to main sequence stars was not anticipated by the pre-discovery theories of planetary formation (see Section 1.3). Such planets were termed “hot Jupiters” due to the expected heating of the planets by the host stars and they were found to have a period cut off at the low end of the period distribution of  $\sim 3$  days. It was only very recently that planets with a period of less than  $\sim 3$  days were discovered: OGLE-TR-56b (Konacki et al. 2003b), OGLE-TR-113b and OGLE-TR-132b (Bouchy et al. 2004). Such planets have been termed “very hot Jupiters”.

In this section, we will consider the subsample of extra-solar planets termed “hot Jupiters” defined by  $P \leq 10.0$  days and excluding the “very hot Jupiters”. The upper limit to the period is chosen to coincide with the start of the “period valley” discussed in Section 1.2.1 but otherwise it is arbitrary. We will try to derive the simplest continuous underlying PDFs for the period  $P$ , orbital semi-major axis  $a_p$ , eccentricity  $e$  and mean mass  $\langle M_p \rangle$ . We will use these analytic representations of the hot Jupiter distribution functions in Chapter 5 when doing Monte Carlo simulations. The analysis below follows the method of Heacox (1999).

Let us define the empirical cumulative distribution function (ECDF) for ordered data  $x_1 \leq x_2 \leq \dots \leq x_N$  by:

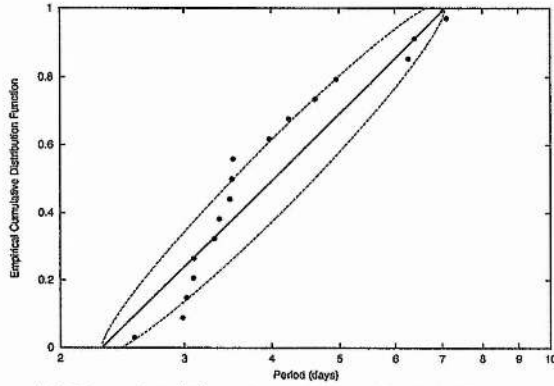
$$F(x_j) = \frac{(j - 0.5)}{N} \quad \text{for } j \in \{1, 2, \dots, N\} \quad (1.1)$$

We have fitted a continuous function  $G(x)$  to each ECDF by iterating a weighted least squares fit with weights  $1/\sigma_{F(x_j)}^2$  (Stuart & Ord 1987) defined by:

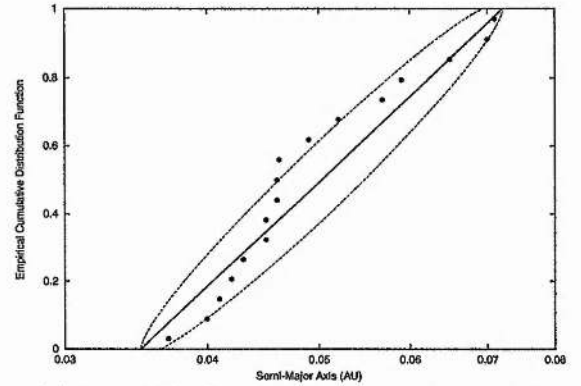
$$\sigma_{F(x_j)}^2 = \frac{G(x_j)(1 - G(x_j))}{N} \quad (1.2)$$

Figure 1.5 shows plots of  $F(x_j)$  against  $x_j$  for each of  $P$ ,  $a_p$ ,  $e$  and  $\langle M_p \rangle$ . Each plot in Figure 1.5 shows the specific function fitted to the ECDF as a continuous line and the  $\pm 1\sigma$  curves as dashed lines. The results of the fits are shown in Table 1.2.

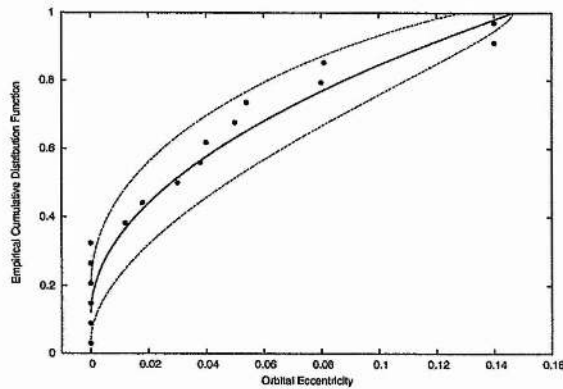
Analysis of the statistical significance of these fits has been done using the Kolmogorov-Smirnov (K-S) goodness of fit test. The K-S test is the most appropriate test in this case



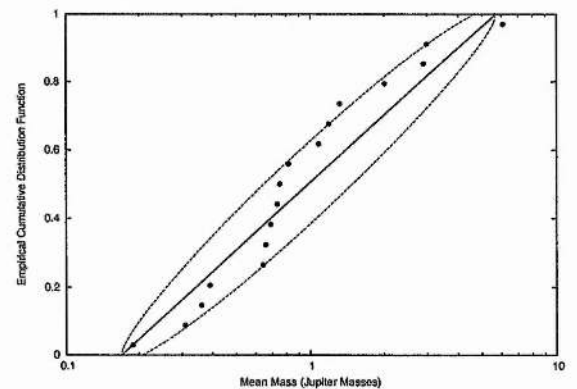
(a) Plot of  $F(P)$  against  $P$  for the hot Jupiters along with the fitted function  $G(P)$ .



(b) Plot of  $F(a_p)$  against  $a_p$  for the hot Jupiters along with the fitted function  $G(a_p)$ .



(c) Plot of  $F(e)$  against  $e$  for the hot Jupiters along with the fitted function  $G(e)$ .



(d) Plot of  $F(\langle M_p \rangle)$  against  $\langle M_p \rangle$  for the hot Jupiters along with the fitted function  $G(\langle M_p \rangle)$ .

Figure 1.5: ECDF plots for each of  $P$ ,  $a_p$ ,  $e$  and  $\langle M_p \rangle$ .

Table 1.2: Results of the fits of  $G(x)$  to  $F(x)$  including the K-S statistic (Col. 7) and its significance (Col. 8).

$x$	Units Of $x$	$G(x)$	$u$	$\sigma_u$	$v$	$\sigma_v$	$k = D_{17}$	$P(D_N \geq k)$
$P$	d	$u + v \log(x)$	-0.734	0.088	2.04	0.13	0.178	0.608
$a_p$	AU	$u + v \log(x)$	4.63	0.23	3.18	0.18	0.176	0.626
$e$	—	$u + v\sqrt{x}$	0.112	0.027	2.32	0.12	0.211	0.389
$\langle M_p \rangle$	$M_J$	$u + v \log(x)$	0.505	0.020	0.655	0.035	0.151	0.797

(Heacox 1999) since we have individual data points (not grouped data). The K-S statistic  $D_N$  is defined by:

$$D_N = \max_{j \in \{1, 2, \dots, N\}} |G(x_j) - F(x_j)| \quad (1.3)$$

Assuming a null-hypothesis that  $x_j$  is drawn from the underlying CDF  $G(x)$  allows one to calculate, for a given  $N$ , the probability of obtaining a value of  $D_N$  greater than or equal to  $k$  where  $k \in \mathbb{R}$ . This probability is denoted by  $P(D_N \geq k)$ . Table 1.2 reports the values of  $k = D_{17}$  obtained for the fits along with the probabilities  $P(D_N \geq k)$  for  $N = 17$ . One can see from Table 1.2 that we cannot reject the null-hypothesis at the  $\alpha = 0.05$  significance level for any of the fits. Hence, for any of  $P$ ,  $a_p$ ,  $e$  and  $\langle M_p \rangle$ , we do not need to hypothesise a closer fit to the data. Summarising the results, we have CDFs of the form:

$$\begin{aligned} G(P) &= u + v \log(P) \\ G(a_p) &= u + v \log(a_p) \\ G(e) &= u + v \sqrt{e} \\ G(\langle M_p \rangle) &= u + v \log(\langle M_p \rangle) \end{aligned} \quad (1.4)$$

The corresponding PDFs are found by differentiating:

$$\begin{aligned} g(P) &\propto P^{-1} \\ g(a_p) &\propto a_p^{-1} \\ g(e) &\propto e^{-0.5} \\ g(\langle M_p \rangle) &\propto \langle M_p \rangle^{-1} \end{aligned} \quad (1.5)$$

#### 1.2.4 Planet Frequency And Stellar Metallicity

The question “What fraction of Sun-like stars have planets, and how does it depend on the host star properties?” is a very important one for planetary formation theories in that the answer may be compared to the theoretical predictions of various models, allowing us to

decide which model is most likely to be correct. As humans, we also want to know if the factors that are responsible for our origin and existence have selected a non-typical location.

The data that we have on extra-solar planets has already shown that Jupiter is a typical giant planet in the sense that it lies in the most densely occupied region of the  $\log(M_p) - \log(P)$  plane (Lineweaver & Grether 2002; Lineweaver, Grether, & Hidas 2003). Analysis of the  $\sim 1800$  Sun-like stars that were being monitored by RV surveys at the end of 2003 has revealed the following facts (Lineweaver & Grether 2003):

1. At least  $\sim 5\%$  of target stars possess at least one planet.
2. Limiting the sample to target stars that have been monitored for  $\gtrsim 15$  years indicates that at least  $\sim 11\%$  possess at least one planet.
3. Limiting the sample to target stars that have been monitored for  $\gtrsim 15$  years and whose low surface activity allows the most precise RV measurements indicates that at least  $\sim 25\%$  possess at least one planet.

It is clear that the longer we survey stars with the RV technique, and the better the accuracy that we achieve, the greater the fraction of stars found to harbour at least one planet. For the work in this thesis, the most pertinent question is “What fraction of main sequence stars have hot Jupiters, and how does this depend on the host star properties?”. Butler et al. (2000) provide an estimate that  $\sim 1\%$  of nearby Sun-like stars (late F and G dwarfs) host a hot Jupiter ( $3.0 \leq P \leq 5.0$  days).

Recently it has come to light that stars with planets seem to be particularly metal rich when compared with “single” (non-binary and no planet found to date) field dwarfs (Santos et al. 2003; Santos, Israelian, & Mayor 2004). The metallicity of a star, denoted by  $[\text{Fe}/\text{H}]$  and with units dex, is defined relative to the Solar metallicity by:

$$[\text{Fe}/\text{H}] = \log(N_{\text{Fe}}/N_{\text{H}})_* - \log(N_{\text{Fe}}/N_{\text{H}})_{\odot} \quad (1.6)$$

where  $\log(N_{\text{Fe}}/N_{\text{H}})_*$  is the log of the ratio of iron abundance to hydrogen abundance for the star, and  $\log(N_{\text{Fe}}/N_{\text{H}})_{\odot}$  is the same ratio for the Sun. Figure 1.6, taken from Santos, Israelian, & Mayor (2004), shows on the left a normalised histogram of the metallicities for two

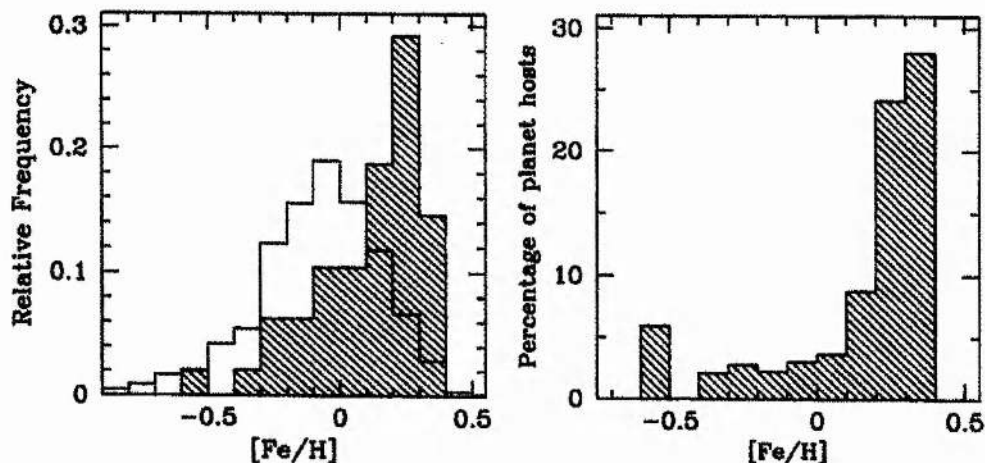


Figure 1.6: Taken from Santos, Israelian, & Mayor (2004). **Left:** Normalised histogram of the stellar metallicities for planet host stars (shaded) and for “single” stars (clear). **Right:** The percentage of stars with planets against stellar metallicity.

star subsamples taken from the volume limited CORALIE star sample (Udry et al. 2000). The shaded histogram represents the metallicity distribution of the 48 stars with known planets, and the clear histogram represents the metallicity distribution of 875 non-binary stars with no planet found to date. One can clearly see the trend to higher metallicities for stars with planets compared to those stars without. Although some authors have claimed that the source of the metallicity excess for planet host stars lies in the process of planetary formation via accretion of disk material and/or planets themselves (Gonzalez 1998), the strongest evidence points to a primordial origin (Sadakane et al. 2002; Santos et al. 2003).

The right hand diagram in Figure 1.6 shows how the percentage of stars with planets depends on stellar metallicity. This diagram clearly demonstrates that the probability of finding a planet is a strong function of the host star metallicity. Approximately  $\sim 28\%$  of stars with  $0.3 \leq [Fe/H] \leq 0.4$  dex host a planet in the CORALIE sample compared to only  $\sim 3\%$  of stars with Solar metallicity ( $[Fe/H] \sim 0.0$  dex) hosting a planet. This result will clearly be of influence in maximising the planet yield of any survey, although choosing high metallicity target stars will introduce a strong bias in the statistics of any planets that are

found.

### 1.2.5 The Variety Of Extra-Solar Planets

At this point it would be appropriate to highlight the variety of extra-solar planetary systems detected to date with a couple of examples. We have already met the hot Jupiters (and very hot Jupiters), and also the pulsar planets, both of whose discovery surprised the astronomical community. It is interesting to note that we know of 8 multiple planetary systems to date (see Table 1.1). In particular the star  $\nu$  Andromedae has a triple planetary system, and it was the first reported multiple planetary system around a main sequence star (Butler et al. 1999). The system consists of a hot Jupiter and two planets more massive than Jupiter in eccentric orbits with semi-major axes 0.83AU and 2.5AU. This is obviously not a Solar System analogue, but it illustrates the possibility that stars with a single extra-solar planet may have other planetary companions with periods too long or masses too small to have been detected yet. Planetary systems are not limited to single stars either. The stellar system of 16 Cygni is actually a triple star system composed of a wide visual binary of two G dwarfs and a distant M dwarf. The G2.5V star 16 Cygni B has a Jupiter mass planetary companion in a very eccentric orbit of semi-major axis 1.7AU (Cochran et al. 1997). Such a variety of planetary systems (in which we must include our own) challenges our theories planetary formation and helps to constrain the likely formation scenarios. Thus it is imperative to keep expanding our database of known extra-solar planets over a range of stellar types and environments.

## 1.3 Star And Planetary Formation

### 1.3.1 Star Formation

Stars are thought to form from gravitational instabilities in interstellar clouds of gas and “dust” grains which lead to collapse and fragmentation (Shu, Adams, & Lizano 1987; Boss 1987). A gravitational instability occurs when the gravitational binding energy of a certain region in the interstellar cloud exceeds the thermal energy (called the Jeans instability criterion). Such an instability may be caused by a shockwave from a supernova for example.



The details of the collapse, including the effects of stellar rotation and magnetic fields, are complex and incompletely known, but they may be roughly split into two stages.

The first stage is characterised by free-fall collapse due to the fact that the inner parts of the cloud contract under self-gravity faster than the outer parts of the cloud. The free-fall time for a test particle of mass  $m$  at the edge of a uniformly collapsing spherical cloud of mass  $M$ , radius  $R$  and initial density  $\rho_0$  may be calculated by considering that it will follow an orbit of semi-major axis  $a = R/2$ , period  $P$  and eccentricity  $e = 1$ . Substituting the expression for the density of the cloud:

$$M = \frac{4\pi R^3 \rho_0}{3} \quad (1.7)$$

into Kepler's third law:

$$P^2 = \frac{4\pi^2 a^3}{GM} \quad (1.8)$$

and using the fact that  $a = R/2$  gives:

$$P = \sqrt{\frac{3\pi}{8G\rho_0}} \quad (1.9)$$

The free-fall time  $t_{\text{ff}}$  is half of the period  $P$  and hence for a cloud of typical density  $10^{-15} \text{kg/m}^3$  the time scale for initial collapse is  $t_{\text{ff}} \sim 10^5$  years.

Interstellar clouds rotate at least a bit, and an isolated rotating cloud must conserve angular momentum. Hence, as the cloud collapses and each particle moves closer to the axis of rotation, the cloud starts to rotate faster. At some point for each particle, the angular speed will become high enough that the centripetal acceleration balances the gravitational force and the particle stops moving closer to the axis of rotation. However, each particle will continue moving parallel to the axis of rotation and as a result, most of the infalling material misses the protostar and ends up in the equatorial plane. This leads to the formation of a disk structure from the initial "spherical" configuration of the interstellar cloud.

The next stage involves the accretion of gas and dust from the disk onto the central object (due to the gravity of the object itself) which heats the centrally condensed gas by compression until nuclear fusion occurs. Infall from the surrounding interstellar cloud replenishes the disk material. The luminescent central object, called a protostar, is formed on timescales of  $10^5$ – $10^6$  years. A hydrodynamic wind from the protostar, channelled by



the geometry of the disk, causes bipolar outflows of gas which carry away  $\sim 10^{-6} - 10^{-4} M_{\odot}$  per year. This mass loss from the system also carries away angular momentum that brakes the rotation of the protostar in  $\sim 10^5$  years.

### 1.3.2 Planet Formation

Planet formation is thought to follow on from the process of star formation via the agglomeration of material within the circumstellar (protoplanetary) disk (Perryman 2000; Ida & Lin 2004). A dust layer starts to form near the central plane of the disk via the sedimentation of dust grains. When the density of the dust layer exceeds some critical value, the dust grains start to stick together as they collide, forming conglomerations called planetesimals. The planetesimals also start colliding with one another, the majority of the collisions being non-destructive, so that the planetesimals start to grow in size. The larger a planetesimal the faster it grows leading to runaway growth. The largest planetesimals sweep up all other planetesimals with orbital semi-major axes similar to their own to become the first protoplanets. The regions near the protoplanets are continuously supplied with planetesimals from the rest of the disk by orbital migration due to gravitational scattering and viscous drag from the disk gas. When a protoplanet reaches a mass of  $\sim 10 M_{\oplus}$  (where  $M_{\oplus}$  is the mass of the Earth), it will start to accrete gas in addition to planetesimals. As this process of planet formation is continuing, the gas and dust of the disk is being blown gradually out of the system by the stellar wind emanating from the protostar until eventually there is very little left to be accreted by the protoplanets. With the supply of planetesimals running out as well, the growth of the protoplanets finally ceases. Table 1.3 shows the approximate timescales and particle/object sizes for each of the stages described above. It should be noted that the timescales shown vary significantly with orbital radius and that the exact details of the formation theory are poorly understood, and so these values should be treated with caution.

The core accretion model described above predicts that the gas giants will form further out in the protosolar system and the terrestrial planets closer in, just as in our own Solar System. This is explained by the following:

Table 1.3: Approximate duration and typical particle/object sizes for each of the stages in the core accretion model of planetary formation (Zeilik & Gregory 1998; Perryman 2000).

Stage	Typical Duration	Typical Particle Size/Mass
Sedimentation of dust grains	$10^4 - 10^5$ years	$0.01\mu\text{m} - 1\mu\text{m}$
Formation of planetesimals	$10^4 - 10^5$ years	1m–1km
Formation of protoplanets	$10^5 - 10^6$ years	$10^{-3}M_{\oplus} - 10^{-1}M_{\oplus}$
Formation of planets	$10^7 - 10^9$ years	$10^{-1}M_{\oplus} - 10M_{\text{J}}$

1. The temperature close to the protostar is too high for gas (and ice) accretion to take place.
2. The gas is blown away first from the inner parts of the protostellar disk.
3. The total mass of the disk material closer in to the protostar is smaller.

As one can see, the core accretion model alone is not sufficient to account for the existence of hot Jupiters. To deal with this, it has become common to invoke orbital migration.

### 1.3.3 Planet Migration

The possibility of planetary migration was actually suggested before the discovery of hot Jupiters (Goldreich & Tremaine 1980). A protoplanet may undergo rapid Type I orbital migration towards the star if it is not yet massive enough to open and sustain a gap in the disk. The migration occurs as a result of torque asymmetries due to the gravitational interaction of the protoplanet with the disk. If the orbital decay time is shorter than the life of the disk, then the protoplanet is in danger of being accreted into the star. At  $\sim 1\text{AU}$  the migration timescale for a  $\sim 1M_{\oplus}$  protoplanet is only  $10^4\text{--}10^5$  years (Ward 1997) compared to a disk lifetime of  $\sim 10^7$  years (Nakano 1987). Also, Type I migration seems inconsistent with the prolific formation of giant extra-solar planets since the protoplanetary cores have a tendency to rapidly migrate to the proximity of their host stars prior to gas accretion.

If gap formation in the disk is successful due to a massive enough protoplanet, then Type II migration will occur. In this case the protoplanet has effectively established a barrier to any radial flow of disk material due to viscous diffusion and the protoplanet becomes locked into the disk (Lin & Papaloizou 1986). Orbital migration is in either direction, and although slower than Type I migration, it may still put a protoplanet in danger of destruction under the right conditions.

The migration scenarios give a possible explanation as to how a gas giant planet may move from the outer parts of a protoplanetary system, where it is most likely to form, to the inner parts of the system. However, without a mechanism for stopping orbital migration,

any protoplanet that migrates to very close distances to the host star ( $\sim 0.05\text{AU}$ ) is most likely to be accreted and we would not observe such an abundance of hot Jupiters. Several potential mechanisms have been put forwards as to how an inwardly migrating planet may be halted at small orbital radii (Lin, Bodenheimer, & Richardson 1996; Trilling et al. 1998). These include the hypothetical existence of a low-density zone maintained by magnetic coupling to the disk, tidal interaction with the spinning star and mass transfer from the protoplanet to the star.

## 1.4 Summary

In this first chapter we have attempted to give the reader an overview of the history of extra-solar planet research, an area of scientific study that has burst to the forefront of the astronomy scene in the space of a decade. The radial velocity technique has dominated the field so far, but this may change in the near future, a possibility we explore in the next chapter. A detailed overview of the characteristics of the extra-solar planets discovered to date has been presented highlighting the finer structure that has not been apparent until recently. We have included our own statistical analysis of the properties of the extra-solar planets as a whole and also concentrating on the class of planets called hot Jupiters, which is of special relevance to the rest of the thesis. We have also considered the frequency of extra-solar planets and the properties of the host stars. Finally we have given a brief summary of star and planetary formation theory, and how the latter has had to invoke migration theory in order to explain the properties of the known extra-solar planets.

# 2

---

## A Menagerie Of Detection Techniques

Up to now, almost all of the extra-solar planet detections have been made by the radial velocity technique despite a number of alternative viable methods being pursued vigorously. This chapter deals with the question “What techniques are being brought to bear on the problem of extra-solar planet detection and what results may we reasonably expect from them?”.

We shall not review every possible detection technique that has been invented (of which there are many). We shall simply review the techniques that have had and/or are expected to have (in the near future) a lot of success. A standard set of parameters is defined below which shall be used throughout, and we assume for simplicity that the star in question has a single planetary companion. For a list of the relevant data/constants that has been used,

see Appendix B.

$L_*, L_p$  – Luminosity of the star/planet.

$M_*, M_p$  – Mass of the star/planet.

$R_*, R_p$  – Radius of the star/planet

$a_*, a_p$  – Semi-major axis of the orbit of the star/planet about the centre of mass.

$P$  – Orbital period of the planet.

$e$  – Orbital eccentricity.

$i$  – Orbital inclination.

$\omega$  – Longitude of periastron.

## 2.1 Radial Velocity

An extra-solar planet and its host star orbit their common centre of mass, and hence the star undergoes periodic variations in its velocity along the line of sight (radial velocity). The semi-amplitude  $K$  of this variation (see Appendix C, Theorem 3) is given by:

$$K = \left( \frac{2\pi G}{P} \right)^{1/3} \frac{M_p \sin i}{(M_* + M_p)^{2/3} (1 - e^2)^{1/2}} \quad (2.1)$$

For a circular orbit ( $e = 0$ ) and for  $M_p \ll M_*$ , the semi-amplitude reduces to:

$$K = \left( \frac{2\pi G}{P} \right)^{1/3} \frac{M_p \sin i}{M_*^{2/3}} \quad (2.2)$$

Due to the influence of Jupiter, the Sun has  $K \approx 12.5 \text{ms}^{-1}$  for  $i = 90.0^\circ$  with a period of 11.9 years. Due to the influence of the Earth, the Sun has  $K \approx 0.09 \text{ms}^{-1}$  for  $i = 90.0^\circ$  with a period of 1.0 years. With the best current precision in radial velocity measurements at  $\sim 2 \text{ms}^{-1}$ , one can see that observations over a period of 12 years are sufficient to detect Jupiter, but there is no chance of detecting the Earth. Also, the  $\sin i$  dependence in Equation 2.1 means that orbital systems seen face on ( $i = 0.0^\circ$ ) result in no radial velocity perturbation.

A radial velocity measurement of a star is made by obtaining a high signal to noise spectrum using a high resolution spectrograph with typically an  $I_2$  gas absorption cell. The  $I_2$  cell provides a wealth of absorption lines superimposed on the stellar spectral lines. This facilitates the measurement of the Doppler shift  $\Delta\lambda$  of the light of wavelength  $\lambda$  arriving from the star relative to the Earth, yielding the star's radial velocity  $v_r = c\Delta\lambda/\lambda$ . Using an accurate ephemeris to correct for the Earth's motion, a set of radial velocity measurements well sampled in time will yield the heliocentric radial velocity curve for the star. The curve may be fitted by the model in Appendix C (Theorem 2), yielding values for  $P$ ,  $e$ ,  $\omega$  and  $M_p \sin i / (M_* + M_p)^{2/3}$ . The value of  $M_*$  may be estimated from the spectral type of the star under observation, and using the approximation  $M_* + M_p \approx M_*$  for  $M_p \ll M_*$ , one may estimate  $M_p \sin i$ , a lower limit for the value of  $M_p$ . To obtain a value of the mean star-exoplanet distance  $a = a_* + a_p \approx a_p$ , one may use Kepler's third law (Equation 1.8 with  $M = M_* + M_p \approx M_*$ ).

The radial velocity technique has been used to discover almost all confirmed extra-solar planets to date making it the most important method so far in the hunt for these objects. In Sections 1.1 and 1.2.5 we have mentioned the relevant references to RV surveys and their discoveries. We have also already noted the  $M_p \sin i$  ambiguity (Appendix A), the intrinsic limit to the precision of RV measurements due to "jitter" and the limitation of a RV survey to the Solar neighbourhood (Section 1.1).

## 2.2 Direct Imaging And Reflected Light

Planets have no intrinsic optical emission. However, a planet will reflect star light from its atmosphere/surface and detecting this star light is a very exciting prospect since it would provide direct and indisputable evidence that extra-solar planetary systems exist. An extra-solar planet will reflect only a small fraction of the incident radiation of wavelength  $\lambda$  in the direction of the line of sight. This fraction depends on the albedo and light scattering properties of the atmosphere/surface, and the phase  $\alpha$  of the planet in its orbit. Writing this fraction as  $p(\lambda, \alpha)$  and considering that the planet is illuminated on one side only, then

we may calculate the luminosity of the extra-solar planet  $L_p$  via reflected light as:

$$L_p = p(\lambda, \alpha) \left( \frac{R_p^2}{2a^2} \right) L_* \quad (2.3)$$

At a distance of 10pc from our Solar System, Jupiter and the Sun have an angular separation of  $\sim 0.52''$  at maximum elongation. In this configuration we would observe half of the planetary disk as illuminated (a half-Jupiter), which leads us to set  $p(\lambda, \alpha) \approx 0.5$  in the best case scenario. Hence Equation 2.3 yields a contrast of  $L_p/L_* \approx 2.1 \times 10^{-9}$ . From the ground, the planetary signal is immersed in the photon noise of the telescope's diffraction profile and the star's seeing profile (typical seeing  $0.5'' - 1.0''$ ) making it undetectable.

Efforts are under way to minimise these problems by employing the following techniques:

1. Suppressing scattered light using a coronagraph.
2. Reduction of the angular size of the star profile using adaptive optics.
3. Observing in the infrared where the value of  $L_p/L_*$  is around  $10^5$  times larger (due to thermal emission from the planet itself).
4. Using interferometric nulling to cancel out the light from the star.
5. Observing from space to eliminate the effects of atmospheric turbulence.

Although the direct imaging method has not been successfully applied to extra-solar planet detection so far, it will in the future be capable of providing broadband colours, spectral features and spectral energy distributions, giving constraints on the temperature and chemical composition of the planet atmosphere/surface, including revealing the presence of biomarkers ( $O_3$ ,  $O_2$ ,  $H_2O$  etc.). This makes direct imaging the technique which has the potential to produce results with the biggest scientific impact.

Individually resolving the star and planet is not absolutely necessary in order to detect the reflected light from the planet. The reflected light from an extra-solar planet is composed of a copy of the star's emission spectrum and an imprint of the planet's absorption spectrum, both red/blue shifted appropriately due to the orbital motion of the extra-solar planet and modulated in strength by the function  $p(\lambda, \alpha)$ . One may obtain a series of



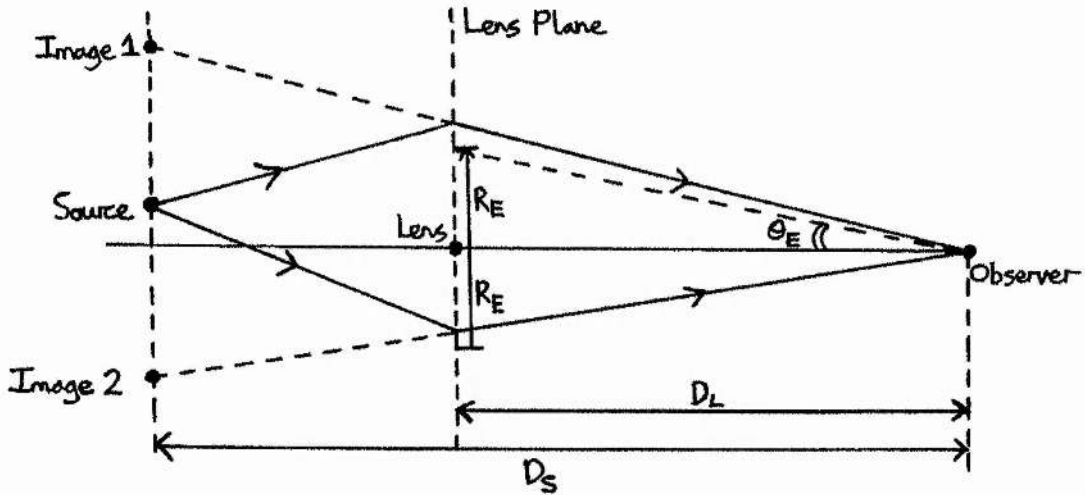


Figure 2.1: General configuration of a gravitational lens.

high resolution spectra and, by carefully modelling and removing the star's spectrum, the reflected light spectrum will be left. A total of three extra-solar planetary systems have been observed in this way (Cameron et al. 2002; Leigh et al. 2003a; Leigh et al. 2003b). Although the reflected spectra were not detected unambiguously, the technique was used to place limits on the geometric albedo  $p$  of each extra-solar planet ( $p < 0.22$  for  $v$  And b,  $p < 0.39$  for  $\tau$  Boo b and  $p < 0.12$  for HD 75289b).

### 2.3 Gravitational Microlensing

Gravitational lensing is the focusing of light rays from a distant source by an intervening massive object (called the lens). The focusing of the light rays produces a magnification of the brightness of the source object. Figure 2.1 shows the configuration of a typical gravitational lens in which the observer sees two images of the source object. If the source lies directly behind the lens as viewed by the observer, then the source forms an image ring at a radius  $R_E$  called the "Einstein radius" of the lens.

Microlensing is the term used to describe gravitational lensing in which the source images as viewed by the observer are too close to be resolved. The Einstein ring radius

(Wambsganss 1997) is given by:

$$R_E = \left[ \frac{4GM_L}{c^2} \frac{(D_S - D_L)D_L}{D_S} \right]^{1/2} \quad (2.4)$$

where  $M_L$  is the mass of the lens object and the distances  $D_S$  and  $D_L$  are as defined in Figure 2.1. The Einstein angle  $\theta_E$  is  $R_E$  expressed in angular units:

$$\theta_E = \frac{R_E}{D_L} \quad (2.5)$$

Looking at source stars in the Galactic bulge ( $D_S \approx 8\text{kpc}$ ) and considering lens stars at a distance  $D_L \approx 4\text{kpc}$  with approximately solar masses ( $M_L \approx 1M_\odot$ ) yields  $\theta_E \approx 0.001''$ . The angular separation of the two source images is approximately  $2\theta_E \approx 0.002''$  which is clearly unresolvable by even the best adaptive optics from the ground (resolution  $\sim 0.1''$ ).

As the source, lens and observer are all in motion, the source will appear to move relative to the lens as viewed by the observer, and the source will undergo a characteristic increase and then decrease in brightness. This is the only observable signature of a microlens, but it is distinguishable from peculiar intrinsic source variability by its achromatic nature. The microlensing magnification  $A(t)$  of the source as a function of time  $t$  is given by:

$$A(t) = \frac{u(t)^2 + 2}{u(t)(u(t)^2 + 4)^{1/2}} \quad (2.6)$$

where  $u(t)$  is the projected distance between the lens and the source in the lens plane in units of  $R_E$  as a function of time  $t$ . Precise alignment of the lens and source is required for a detectable source brightening and hence the probability of substantial microlensing magnification is extremely small ( $\sim 10^{-6}$  for source stars in the Galactic bulge). Timescales for a typical Galactic bulge microlensing event range from days to months (Wambsganss 1997).

The presence of a planet orbiting the lens at a projected distance from the lens close to the Einstein radius ( $0.6-1.6R_E$ ) will produce a secondary amplification of the source, shorter in timescale than the primary event (of the order of a couple of hours for an Earth mass planet and of the order of a day for a Jupiter mass planet), but theoretically of unlimited magnification. Again, assuming  $D_S \approx 8\text{kpc}$  for Galactic bulge stars and considering a lens star at a distance  $D_L \approx 4\text{kpc}$  with  $M_L \approx 1M_\odot$  yields  $R_E \approx 4.0\text{AU}$ , similar to Jupiter's distance from the Sun.

With the advent of observational programmes monitoring millions of stars (OGLE: Udalski et al. 1993; MACHO: Alcock et al. 1993 etc.), hundreds of photometric microlensing events have now been observed. A small subset of the already rare microlensing events have revealed anomalies which are due to a number of different scenarios including binary lenses (Udalski et al. 1994) and the effect of the Earth's motion around the Sun (Alcock et al. 1995). It is only recently that the first unambiguous planetary microlensing event has been detected (Bond et al. 2004) from the perturbation of the simple lightcurve of an isolated point lens.

The main drawback to the planetary microlensing technique is that a microlensing event only occurs once, requiring all relevant observations to be taken during the event. The only follow-up observations that may be possible are those of the lens star in the years following the event due to the proper motion of the lens with respect to the source. However, planetary microlensing has the advantages that it can find planetary systems at large distances (in the kpc regime) and that it is sensitive down to Earth mass planets that lie close to the Einstein ring (which fortuitously coincides with distances of the order of 1 AU). Although this technique may only capture a handful of extra-solar planets in the near future, its main power lies in its ability to acquire statistics on many events. Microlensing surveys of the Galactic bulge have consequently been used to place constraints on the abundance of Galactic planets based on the observed rate (or lack) of planetary microlensing events. From 5 years of PLANET photometry of microlensing events, Gaudi et al. (2002) concluded that less than 33% of M dwarfs in the Galactic bulge have companions with mass  $M_p = 1M_J$  between 1.5 and 4AU, and that less than 45% have companions with mass  $M_p = 3M_J$  between 1 and 7AU. More recently Snodgrass, Horne, & Tsapras (2004) found that the Galactic abundance of  $\sim 1M_J$  planets at  $\sim 4$  AU has an upper limit of  $5n\%$  based on the sample of OGLE-II and OGLE-III microlensing events (where  $n$  is the number of planetary microlensing events detected in this sample and  $n \leq 2$ ).

## 2.4 Transit Photometry

Given a fortuitous geometric alignment, an extra-solar planet may be observed to eclipse the host star as viewed from the Earth. Such an eclipse is called a planetary transit and it is characterised by a small decrease in the observed brightness of the host star that repeats at the orbital period of the extra-solar planet.

### 2.4.1 Transit Probability

Consider an extra-solar planet P of radius  $R_p$  orbiting a star S of radius  $R_*$  in a circular orbit of radius  $a$ . Let  $i$  be the orbital inclination of the planet (the angle between the line of sight and the normal to the orbital plane). Figure 2.2 shows the configuration of S and P when the disk of P just touches the disk of S at the “top” of its orbit as projected in the sky plane. It is clear from Figure 2.2 that in order for a transit to occur, the disk of P must obscure the disk of S to some extent. This condition may be written as:

$$x = a \sin(90^\circ - i) < R_* + R_p \quad (2.7)$$

Hence the probability  $P_{\text{tra}}$  of observing a transit is given by:

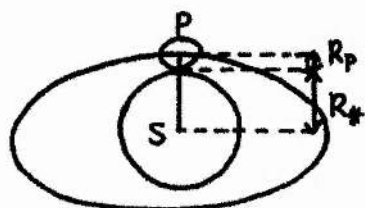
$$\begin{aligned} P_{\text{tra}} &= P(a \sin(90^\circ - i) < R_* + R_p) \\ &= 1 - P\left(i \leq \cos^{-1}\left(\frac{R_* + R_p}{a}\right)\right) \end{aligned} \quad (2.8)$$

Under the assumption that the orbital inclination of the extra-solar planet P is random, we can apply Equation [6] in Appendix A to Equation 2.8, which yields:

$$P_{\text{tra}} = \frac{R_* + R_p}{a} \quad (2.9)$$

At this point it is useful to note the difference between a grazing eclipse and an annular eclipse for an eclipsing body that is smaller than the host star. A grazing eclipse occurs when the disk of the eclipsing body lies only partially in front of the disk of the host star at all times during the eclipse. An annular eclipse occurs when the disk of the eclipsing body

View along the line of sight:



Side view:

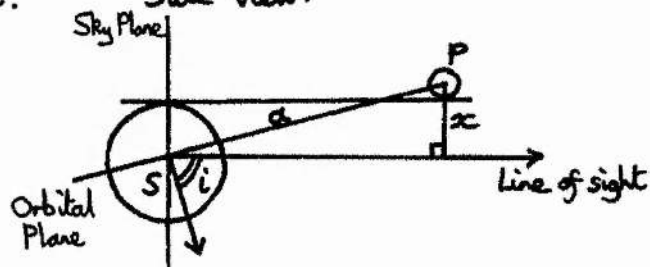
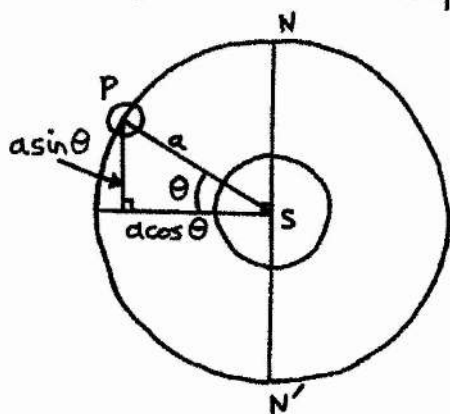


Figure 2.2: Configuration of an extra-solar planet and the host star when the disk of the planet is observed to touch the disk of the star at the "top" of its orbit as projected in the sky plane.

View from above the orbital plane:



View along the line of sight:

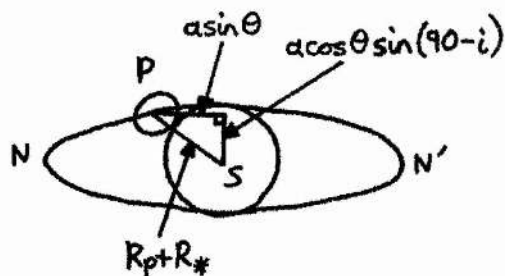


Figure 2.3: General configuration of an extra-solar planet and the host star at the start of the transit ingress.

lies fully in front of the disk of the host star at some point during the eclipse. Following similar arguments to those used above, it may be shown that the probability of an annular eclipse  $P_{\text{ann}}$  of S by P is given by:

$$P_{\text{ann}} = \frac{R_* - R_p}{a} \quad (2.10)$$

Since an eclipse is either annular or grazing we have:

$$P_{\text{tra}} = P_{\text{ann}} + P_{\text{gra}} \quad (2.11)$$

where  $P_{\text{gra}}$  is the probability of a grazing eclipse of S by P. Rearranging Equation 2.11 and using Equations 2.9 and 2.10 yields:

$$P_{\text{gra}} = \frac{2R_p}{a} \quad (2.12)$$

It is clear from Equation 2.9 that the closer a planet is to the host star, the more likely it is to be observed to transit the stellar disk as seen from the Earth. In fact, for a star of a particular radius,  $P_{\text{tra}}$  is maximised for a hot Jupiter. For example, consider a typical hot Jupiter ( $a \approx 0.05\text{AU}$  and  $R_p \approx 1.4R_J$  - see Section 2.4.5) orbiting a Sun like star ( $R_* \approx 1R_\odot$ ). Then we calculate  $P_{\text{tra}} \approx 0.106$ . The strongest transit signal occurs for an annular eclipse with  $P_{\text{ann}} \approx 0.080$  for a hot Jupiter. However, an Earth analogue has  $P_{\text{tra}} \approx 4.70 \times 10^{-3}$  and  $P_{\text{ann}} \approx 4.61 \times 10^{-3}$ , while a Jupiter analogue has  $P_{\text{tra}} \approx 9.86 \times 10^{-4}$  and  $P_{\text{ann}} \approx 8.03 \times 10^{-4}$ . The conclusion drawn from this discussion is that the transit technique favours the detection of hot Jupiters over other types of planets from a purely probabilistic point of view. If a Sun-like star hosts a hot Jupiter, then it has a  $\sim 10\%$  chance of exhibiting a periodic transit signal although this does not imply that we will be able to detect such a signal.

The assumption that an extra-solar planet resides in a circular orbit ( $e \leq 0.1$ ) is not valid for most known extra-solar planets (see Table 1.1). However, the hot Jupiters considered in Section 1.2.3 all have  $e \leq 0.15$  (approximately circular orbits) which is to be expected since this type of planet is likely to have undergone orbital circularisation via tidal interaction with the host star. Hence, the circular orbit assumption is valid for this class of extra-solar planets, as it is for the Earth ( $e = 0.017$ ) and Jupiter ( $e = 0.048$ ). Motivation for the

derivations above based on this assumption lies in the fact that hot Jupiters are the focus of this thesis and more detailed calculations involving eccentric orbits are not required.

### 2.4.2 Transit Duration

In order to derive an expression for the duration  $\Delta t$  of a transit event, we consider the same set up as in Section 2.4.1. The duration of a transit event is defined as the time elapsed from when the disk of P first touches the disk of S (the start of the transit ingress) to when the disk of P last touches the disk of S (the end of the transit egress), all as projected in the sky plane. Figure 2.3 shows the general configuration of S and P at the start of the transit ingress. The line NN' is the line of nodes (see Figure C.1). From the geometry presented in Figure 2.3, we may write the following equation (Pythagoras's Theorem):

$$a^2 \sin^2 \theta + a^2 \cos^2 \theta \sin^2(90 - i) = (R_* + R_p)^2$$

$$\therefore \sin^2 \theta + \cos^2 \theta \cos^2 i = \left( \frac{R_* + R_p}{a} \right)^2 \quad (2.13)$$

$$\therefore \sin^2 \theta \sin^2 i + \cos^2 i = \left( \frac{R_* + R_p}{a} \right)^2$$

$$\therefore \theta = \sin^{-1} \left[ \left( \frac{1}{\sin i} \right) \sqrt{\left( \frac{R_* + R_p}{a} \right)^2 - \cos^2 i} \right] \quad (2.14)$$

The extra-solar planet P has a constant angular speed  $\omega$  in its orbit given by:

$$\omega = \frac{2\pi}{P} \quad (2.15)$$

where  $P$  is the orbital period. During the duration  $\Delta t$  of the transit event, P moves an angle of  $2\theta$  around its orbit at constant angular speed  $\omega$ . Hence we also have:

$$\omega = \frac{2\theta}{\Delta t} \quad (2.16)$$

Substituting Equation 2.16 into Equation 2.15 and rearranging gives:

$$\Delta t = \frac{P\theta}{\pi} \quad (2.17)$$



Substituting Equation 2.14 into Equation 2.17 yields the following analytic expression for the transit duration  $\Delta t$ :

$$\Delta t = \frac{P}{\pi} \sin^{-1} \left[ \left( \frac{1}{\sin i} \right) \sqrt{\left( \frac{R_* + R_p}{a} \right)^2 - \cos^2 i} \right] \quad (2.18)$$

For all known extra-solar planets (except maybe the very hot Jupiters) we have  $R_* + R_p \ll a$  and hence  $\theta$  is small. In this case we have  $\sin \theta \approx \theta$  and  $\cos \theta \approx 1$ . Using this result in Equation 2.13, rearranging and then using Equation 2.17 leads to the better known expression for  $\Delta t$ :

$$\Delta t = \frac{P}{\pi} \sqrt{\left( \frac{R_* + R_p}{a} \right)^2 - \cos^2 i} \quad (2.19)$$

A central transit occurs when  $i = 90.0^\circ$ . In this case we have a transit duration  $\Delta t_{\text{cen}}$  given by:

$$\begin{aligned} \Delta t_{\text{cen}} &= \frac{P}{\pi} \sin^{-1} \left( \frac{R_* + R_p}{a} \right) \\ &\approx \frac{P(R_* + R_p)}{\pi a} \end{aligned} \quad (2.20)$$

where the approximation is valid for  $R_* + R_p \ll a$ .

Let us consider our examples from Section 2.4.1 again. We may calculate  $P$  from Equation 1.8 using the assumed value for  $a$  and noting that  $M \approx 1M_\odot$ . Then, using Equation 2.20, we calculate that  $\Delta t_{\text{cen}} \approx 3.33$  hours for the hot Jupiter,  $\Delta t_{\text{cen}} \approx 13.1$  hours for the Earth and  $\Delta t_{\text{cen}} \approx 32.6$  hours for Jupiter. A typical night of photometric observations lasts  $\sim 8$  hours from the ground. Consequently, the hot Jupiter transit signal can fit into a single night of observations whereas an Earth or Jupiter transit signal will only be detectable by comparing different nights of observations, making them more difficult to find, especially if the atmospheric conditions change considerably from one night to the next.

### 2.4.3 Transit Lightcurve Morphology

The transit of an extra-solar planet across the disk of the host star as viewed from the Earth temporarily obscures a fraction of the luminous stellar disk. Since we cannot resolve the



stellar disk from the Earth (the star is well modelled by a point source of light), we simply observe a small temporary decrease in the brightness of the host star during a transit event. In order to calculate the predicted lightcurve of an extra-solar planetary transit event, we have adopted the following model.

A luminous and spherical primary star was assumed with an orbiting dark companion (planets have no intrinsic optical emission and reflect very little light - see Section 2.2). The surface brightness of the observed disk of the star was assumed to obey a linear limb darkening law:

$$I(\mu) = I(1)(1 - u(1 - \mu)) \quad (2.21)$$

where  $\mu$  is the cosine of the angle between the normal to the stellar surface and the line of sight,  $I(1)$  is the surface brightness at the centre of the stellar disk,  $I(\mu)$  is the surface brightness of the stellar disk as a function of  $\mu$  and  $u$  is the linear limb darkening coefficient ( $0 \leq u \leq 1$ ). The companion was assumed to be spherical and massless (the most massive extra-solar planets have  $M_p \sim 10M_J \approx 9.5 \times 10^{-3}M_\odot$ ). The orbit of the companion was assumed to be circular (true for hot Jupiters and most planets in our Solar System) with the star fixed at the centre of the orbit. The observed stellar flux  $f(t)$  at time  $t$  is then given by:

$$f(t) = f_0(1 - f_c(t)) \quad (2.22)$$

where  $f_0$  is the total stellar flux when the star is unobscured and  $f_c(t)$  is the fraction of the total stellar flux obscured by the companion at time  $t$ .

We have developed a function called *transitcurve.pro* in the programming language IDL that calculates the function  $f(t)/f_0$  in a numerical fashion. It works by making a grid for the orbital phase  $t/P$  over the range  $-0.5 \leq t/P \leq 0.5$ , where  $P$  is the orbital period. We define  $t/P = 0$  when the companion is at its closest to the observer. For each orbital phase  $t/P$ , *transitcurve.pro* creates a grid for the observed stellar disk and calculates the flux from each grid element taking into account the apparent position of the companion and the effects of linear limb darkening. The sum of the fluxes from all the grid elements yields the value of  $f(t)$  at phase  $t/P$  which is then normalised by  $f_0$ .

The function *transitcurve.pro* requires the arguments  $R_p/R_*$ ,  $a/R_*$ ,  $i$  (degrees) and  $u$

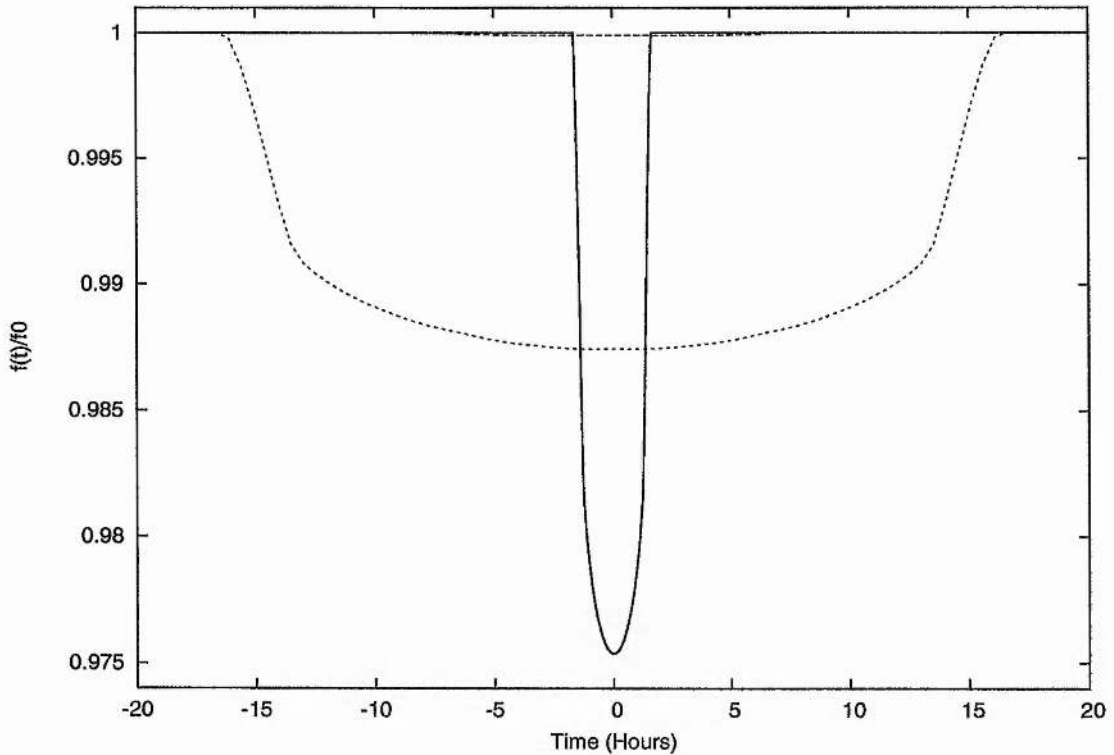


Figure 2.4: Transit lightcurves for a typical hot Jupiter (continuous line), the Earth (dashed line) and Jupiter (shorter dashed line) orbiting the Sun ( $M_* = 1M_\odot$  and  $R_* = 1R_\odot$ ) with  $i = 90.0^\circ$  and  $u = 0.5$ .

in order to define the physical situation. It also requires integer values for the stellar grid resolution  $G$  and the lightcurve phase resolution  $L$ . The function *transitcurve.pro* creates a grid for the stellar disk of radius  $G$  grid elements inside a square of side  $2G$  grid elements, and it returns a lightcurve with  $L$  data points. Each lightcurve data point contains the value of the time  $t$  in units of the orbital period (orbital phase) and the corresponding value of  $f(t)/f_0$ . We have made two adjustments to improve the speed of *transitcurve.pro*. Firstly we note that the transit curve is symmetric about the phase  $t/P = 0$ , and so we only carry out the stellar grid calculations for the orbital phase range  $-0.5 \leq t/P \leq 0.0$ . Secondly, we limit the stellar grid calculations to a square around the companion at each orbital phase  $t/P$  and calculate by summation the total obscured stellar flux instead of the total observed stellar flux.

Let us define the depth  $\Delta f/f_0$  of a transit event by:

$$\frac{\Delta f}{f_0} = \max_t(f_c(t)) \quad (2.23)$$

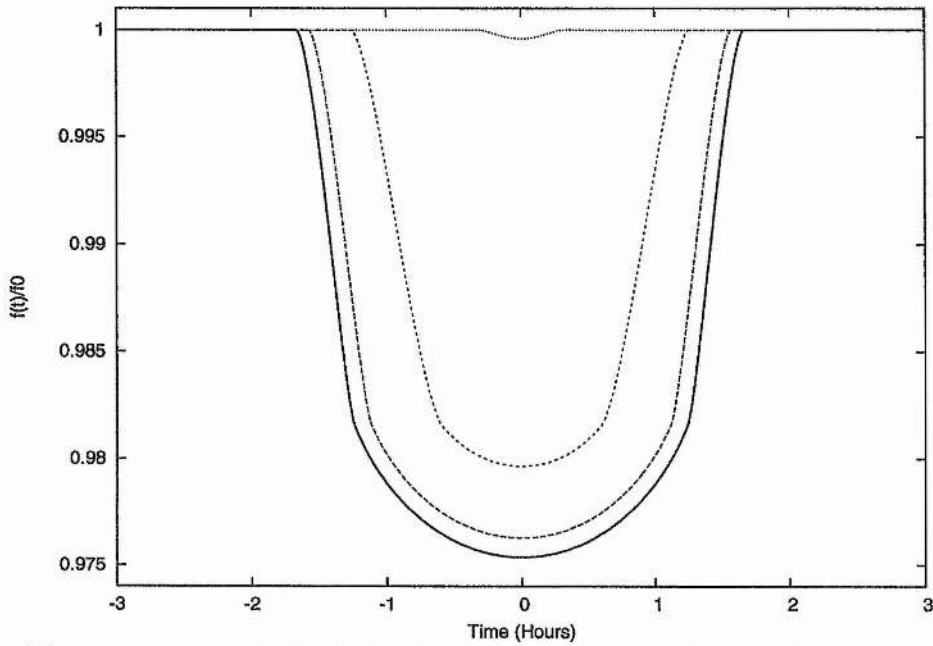
Figure 2.4 shows the theoretical transit lightcurves calculated by *transitcurve.pro* for a typical hot Jupiter (as defined in Section 2.4.1 and represented by a continuous line), the Earth (dashed line) and Jupiter (shorter dashed line) orbiting the Sun ( $M_* = 1M_\odot$  and  $R_* = 1R_\odot$ ) with  $i = 90.0^\circ$  and  $u = 0.5$ . It can be seen that the hot Jupiter causes the deepest transit depth ( $\Delta f/f_0 \approx 2.5\%$ ) and that Jupiter causes a transit depth of  $\Delta f/f_0 \approx 1.3\%$  while the Earth causes a minute transit depth of  $\Delta f/f_0 \approx 0.01\%$ . Figure 2.5(a) shows the effect of varying orbital inclination on the transit lightcurve for the hot Jupiter orbiting the Sun (with  $u = 0.5$ ). The continuous, dashed, shorter dashed and dotted lightcurves correspond to inclinations of  $90.0^\circ$ ,  $88.0^\circ$ ,  $86.0^\circ$  and  $84.0^\circ$  respectively. One can see that as the inclination of the orbit of an extra-solar planet decreases from  $90.0^\circ$ , the transit depth and duration decrease until, at a certain inclination, there is no eclipse. Figure 2.5(b) shows the effect of varying values for the linear limb darkening coefficient on the transit lightcurve for the hot Jupiter orbiting the Sun (with  $i = 90.0^\circ$ ). The dotted, shorter dashed, continuous and dashed lightcurves correspond to values of the linear limb darkening coefficient given by 0.0, 0.25, 0.5 and 1.0 respectively. One can see that when there is no limb darkening ( $u = 0.0$ ), the transit profile is flat bottomed. As  $u$  increases from zero to its maximum value of  $u = 1.0$ , the transit curve becomes smoother and deeper. Note that the increase in  $\Delta f/f_0$  with increasing  $u$  is only true for orbital inclinations close to  $90.0^\circ$ .

If we ignore the effects of limb darkening, then we can estimate the transit depth  $\Delta f/f_0$  during an annular eclipse as follows. With no limb darkening ( $u = 0$ ), the stellar disk (radius  $R_*$ ) has a uniform surface brightness  $I$  (from Equation 2.21). The total stellar flux  $f_0$  is then given by:

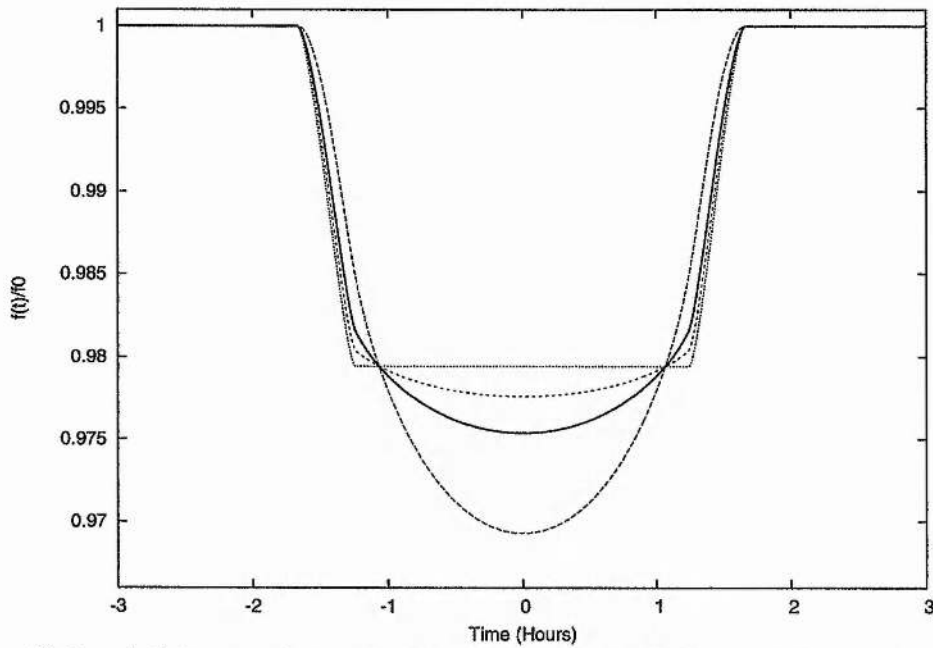
$$f_0 = I\pi R_*^2 \quad (2.24)$$

During an annular eclipse by a planet with radius  $R_p$ , an area  $\pi R_p^2$  of the stellar surface is obscured, which corresponds to a flux  $\Delta f$  given by:

$$\Delta f = I\pi R_p^2 \quad (2.25)$$



(a) Transit lightcurves for a typical hot Jupiter orbiting the Sun for different values of  $i$  with  $u = 0.5$ . The continuous, dashed, shorter dashed and dotted lightcurves correspond to inclinations of  $90.0^\circ$ ,  $88.0^\circ$ ,  $86.0^\circ$  and  $84.0^\circ$  respectively.



(b) Transit lightcurves for a typical hot Jupiter orbiting the Sun for different values of  $u$  with  $i = 90.0^\circ$ . The dotted, shorter dashed, continuous and dashed lightcurves correspond to values of the linear limb darkening coefficient given by 0.0, 0.25, 0.5 and 1.0 respectively.

Figure 2.5: Transit lightcurve morphology as a function of inclination and the linear limb darkening coefficient.

Dividing Equation 2.25 by Equation 2.24 yields:

$$\frac{\Delta f}{f_0} = \left( \frac{R_p}{R_*} \right)^2 \quad (2.26)$$

Equation 2.26 yields  $\Delta f/f_0 \approx 0.0206$  for the hot Jupiter,  $\Delta f/f_0 \approx 8.4 \times 10^{-5}$  for the Earth and  $\Delta f/f_0 \approx 0.0105$  for Jupiter, in rough agreement with Figure 2.4. It is clear from Equation 2.26 that the transit depth  $\Delta f/f_0$ , and therefore the strength of transit signal, is heavily dependent on the relative sizes of the star and extra-solar planet. The transit depth is maximised for large planets (Jupiter size) orbiting small stars (M dwarfs for example).

Regular photometric observations (many per transit duration) of the host star of sufficient accuracy (better than the transit depth) are required in order to define a transit lightcurve. Fitting the model described above (and encapsulated in Equation 2.22) to the data in flux units allows one to determine the orbital period  $P$  (from observations of multiple transits), the time of mid-transit  $t_0$ , the “out of transit” magnitude of the star  $m_0$  (where  $m_0 = -2.5 \log(f_0)$ ), the planet to star radius ratio  $R_p/R_*$  and the impact parameter  $b$  (where  $b = a \cos i/R_*$ ). The values of  $M_*$  and  $R_*$  may be estimated from the spectral type and/or broad-band colours of the star under observation. With this information, one can calculate the value of  $a$  from Kepler’s third law (Equation 1.8), and fit a value for the orbital inclination  $i$  instead of the impact parameter  $b$ .

#### 2.4.4 Transit Signal To Noise

In the previous section we derived an approximation for the transit depth  $\Delta f/f_0$  of a transiting extra-solar planet. We will now derive an expression for the transit signal to noise S/N based on some simple assumptions and then use this to determine what sort of transits we may be able to detect from ground based observations.

Consider a box-car transit lightcurve with depth  $\Delta f/f_0$  that has  $N_{\text{in}}$  photometric observations during transit and  $N_{\text{out}}$  photometric observations out of transit. Let each photometric observation have an uncertainty  $\sigma f_0$  (see Figure 2.6). The signal  $S$  of the transit

curve is defined by:

$$S = f_0 - \min_t(f(t)) \quad (2.27)$$

$$= f_0 - f_{\text{in}} \quad (2.28)$$

where  $f_{\text{in}}$  is the constant observed stellar flux during transit (since the transit curve is a box-car). Substituting Equation 2.22 into Equation 2.27 we get:

$$\begin{aligned} S &= f_0 - f_0(1 - \max_t(f_c(t))) \\ &= \Delta f \end{aligned} \quad (2.29)$$

By considering Equation 2.28, the uncertainty in  $S$ , denoted by  $\sigma_S$ , may be written as:

$$\sigma_S^2 = \sigma_{f_0}^2 + \sigma_{f_{\text{in}}}^2 \quad (2.30)$$

where  $\sigma_{f_0}$  is the uncertainty in  $f_0$  and  $\sigma_{f_{\text{in}}}$  is the uncertainty in  $f_{\text{in}}$ . The values of  $f_0$  and  $f_{\text{in}}$  may be estimated from the photometric data by calculating the mean of the photometric data out of transit and the mean of the photometric data during transit respectively. Hence the uncertainties in  $f_0$  and  $f_{\text{in}}$  are given by:

$$\sigma_{f_0}^2 = \frac{\sigma^2 f_0^2}{N_{\text{out}}} \quad (2.31)$$

$$\sigma_{f_{\text{in}}}^2 = \frac{\sigma^2 f_0^2}{N_{\text{in}}} \quad (2.32)$$

Substitute Equations 2.31 and 2.32 into Equation 2.30:

$$\sigma_S^2 = \frac{\sigma^2 f_0^2}{N_{\text{out}}} + \frac{\sigma^2 f_0^2}{N_{\text{in}}} \quad (2.33)$$

The signal to noise S/N of the transit curve is given by:

$$S/N = \frac{S}{\sigma_S} \quad (2.34)$$

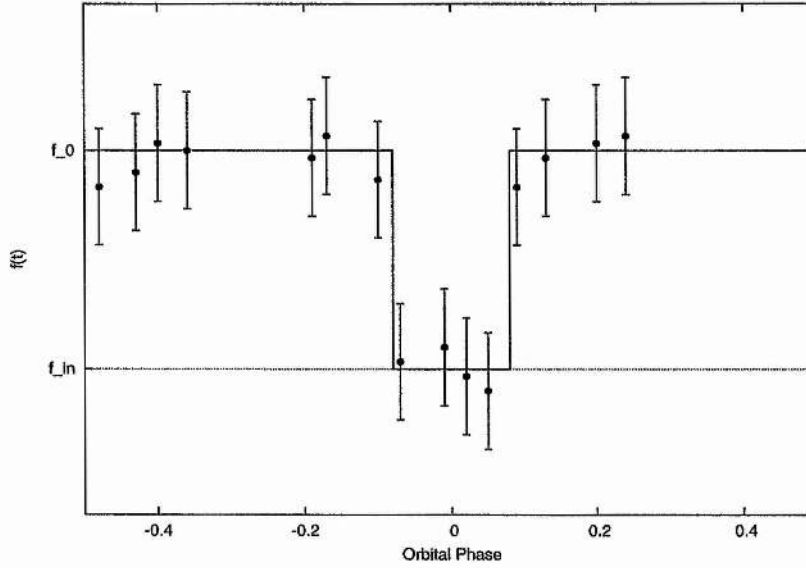


Figure 2.6: Typical photometric observations of a transit event modelled by a box-car transit lightcurve

Substitute Equations 2.29 and 2.33 into Equation 2.34:

$$S/N = \frac{\Delta f}{\sigma f_0} \sqrt{\frac{N_{out} N_{in}}{N_{out} + N_{in}}} \quad (2.35)$$

This expression for the signal to noise shows that it is important to obtain sufficient photometric observations not just during the transit event but also out of transit. To illustrate this point, consider Equation 2.35 in the limit  $N_{in} \gg N_{out}$ . Then  $S/N \propto \sqrt{N_{out}}$ , and the number of observations out of transit  $N_{out}$  limits the S/N achieved. Similarly, in the limit  $N_{out} \gg N_{in}$ , we have  $S/N \propto \sqrt{N_{in}}$ , and the number of observations during transit  $N_{in}$  limits the S/N achieved. In general it is easier to have  $N_{out} \gg N_{in}$  due to the fact that a transit event lasts a small fraction of the orbital period, and most observations will therefore be out of transit.

A firm detection of a transit candidate requires at least a  $4\sigma$  confidence (or equivalently a  $S/N \geq 4$ ). However, a  $6\sigma$  detection threshold is generally required in order to limit the number of false alarms. The best photometry from the ground currently achieves an accuracy of  $\sim 0.2 - 0.3\%$ , and regularly achieves an accuracy of  $\sim 1\%$  for the brighter stars in the star sample. The accuracy is limited by photon noise (mainly from the sky background)



and complicated by varying extinction, seeing etc. Let us assume that photometric observations with an uncertainty of  $\sigma \approx 1.0\%$  can be made of a transiting extra-solar planetary system every 5 minutes during a typical observing run that lasts 10 nights with 8 hours of observations per night (960 lightcurve data points). Lets also assume that a central transit is observed in the middle of night four (arbitrary). Considering the examples from Section 2.4.1 again, we have for the hot Jupiter  $P \approx 4.08$  days and  $\Delta f/f_0 \approx 0.0206$  ignoring the effects of limb darkening. Hence 2 transits will be observed, and since  $\Delta t_{\text{cen}} \approx 3.33$  hours, we have  $N_{\text{in}} \approx 80$  and  $N_{\text{out}} \approx 880$ . Equation 2.35 then yields a S/N of 17.6, well above the detection threshold. For the Earth, we have  $P = 1$  year and  $\Delta f/f_0 \approx 8.4 \times 10^{-5}$ . Hence 1 transit is observed of duration  $\Delta t_{\text{cen}} \approx 13.1$  hours. However, only 8 hours during the transit are observed and consequently we have  $N_{\text{in}} \approx 96$  and  $N_{\text{out}} \approx 864$ . The S/N achieved is 0.078, clearly a non-detection. Finally, for Jupiter, 8 hours of 1 transit are observed with  $\Delta f/f_0 \approx 0.0105$ . Therefore  $N_{\text{in}} \approx 96$  and  $N_{\text{out}} \approx 864$  and the S/N achieved is 9.76, more than sufficient for a detection.

It is also interesting to calculate the transit depth  $\Delta f/f_0$  corresponding to the detection threshold of S/N = 4 given the best current photometric observations ( $\sigma \approx 0.2\%$ ) with  $N_{\text{in}} \approx 80$  and  $N_{\text{out}} \approx 880$  (based on the typical hot Jupiter). The transit depth corresponding to a S/N of 4 is then  $\Delta f/f_0 \approx 9.34 \times 10^{-4}$ , which, for a Sun like star ( $R_* \approx 1R_{\odot}$ ), corresponds to a planetary radius of  $R_p \approx 0.30R_J$ . This is a good estimate of the current limit from the ground to the size of an extra-solar planet detectable by the transit technique, although theoretically it may be improved by higher cadence observations and by observing smaller stars.

#### 2.4.5 The History Of The Transit Technique

Five years after the discovery of 51 Peg b, the extra-solar planet in orbit around HD 209458 was found to transit the stellar disk (Charbonneau et al. 2000). This hot Jupiter was already known to have  $M_p \sin i = 0.69 \pm 0.05 M_J$  from the RV measurements (Mazeh et al. 2000). Also, the spectral type, and hence the mass and radius, of the host star were already known. Consequently, the modelling of the two observed transit events allowed the measurement of



the orbital inclination, which in turn allowed the true mass of HD 209458b to be calculated. Charbonneau et al. (2000) measured  $i = 87.1 \pm 0.2^\circ$  implying that  $M_p = 0.69 \pm 0.05 M_J$ . They also measured  $R_p = 1.27 \pm 0.02 R_J$  from the transit fit.

The importance of this result lies in the fact that for the first time the mass and radius of an extra-solar planet had been measured, not just a lower limit on the mass. Before this discovery, the radii of the extra-solar planets were unknown and hence their average densities were also unknown. The average density derived for HD 209458b was  $\sim 0.38 \text{ g/cm}^3$ , significantly less than the average density of Saturn ( $0.7 \text{ g/cm}^3$ ), the least dense of the Solar System gas giants. This was proof that HD 209458 must be a gas giant rather than a rocky (terrestrial) planet, lending weight to the term hot Jupiter. Average density was not the only important quantity that could be calculated for an extra-solar planet for the first time. Other such quantities included surface gravity and effective temperature.

Other groups soon followed in observing the transits of HD 209458, the most notable being Hubble Space Telescope (HST) observations (Brown et al. 2001) which provided exquisite detail on the morphology of the transit lightcurve with a photometric precision of  $\sim 0.11 \text{ mmag}$ . Brown et al. (2001) were able to estimate the linear limb darkening coefficient of the host star and even put limits on the sizes and masses of any circumplanetary rings or satellites. They inferred a planetary radius of  $\sim 1.35 R_J$ , similar to the value of  $\sim 1.40 R_J$  previously reported by Mazeh et al. (2000). The uncertainty in the mass and radius of the host star is the main source of uncertainty in the derived planetary radius for the above estimates. Cody & Sasselov (2002) used detailed modelling of the host star to derive their estimate for the planetary radius of  $\sim 1.42 \pm 0.13 R_J$ . We adopt the value  $R_p \approx 1.4 R_J$  for the radius of a typical hot Jupiter (see Section 2.4.1).

Since the discovery of the transiting nature of HD 209458 b, many transit candidates have been put forward by various groups. OGLE have been by far the most prolific transit survey with over 100 transit candidates from two observational seasons (Udalski et al. 2002a; Udalski et al. 2002b; Udalski et al. 2003). EXPLORE have produced a handful of transit candidates that are currently being followed up spectroscopically (Mallén-Ornelas et al. 2003) and a search of the MACHO photometry database has revealed nine transit candidates (Drake & Cook 2004). However, only three transit candidates have been con-

firmed as extra-solar planets: OGLE-TR-56b (Konacki et al. 2003b), OGLE-TR-113b and OGLE-TR-132b (Bouchy et al. 2004).

A very interesting result, and one that is relevant to this thesis, is that Gilliland et al. (2000) found a lack of planets in the globular cluster 47 Tuc from HST observations. More specifically they observed  $\sim 34000$  main sequence stars in the cluster over 8.3 days. Assuming a typical hot Jupiter with  $R_p = 1.30R_J$ ,  $P = 3.5$  days and  $P_{\text{tra}} = 0.10$ , and a hot Jupiter fraction of 1.0% (same as the Solar neighbourhood), Gilliland et al. (2000) expected to have detected  $\sim 17$  transiting hot Jupiters. They actually found no transit signals that could be the result of transiting planets using their two-colour time-series photometry (see Section 2.4.6) and hence they could conclude with a very high confidence that the hot Jupiter fraction for 47 Tuc is at least an order of magnitude lower than that for the Solar neighbourhood. The cause of the absence of hot Jupiters in 47 Tuc is not known, but it was suggested that low metallicity and/or crowding in the cluster interfered with planet formation, orbital migration or planet survival.

#### 2.4.6 Transit Mimics And Follow-Up Strategies

Once a transit signal has been detected, it needs to be confirmed as an extra-solar planet via the analysis of appropriate follow-up observations. However there are many scenarios which may result in a similar lightcurve to that of a real transiting planet and any follow-up observations need to be designed carefully in order to rule out these transit “mimics” if possible. Below we describe the possible sources of confusion and what type of follow-up observations may be used to determine the correct scenario.

The first type of transit mimic may come from stellar activity in the form of star spots. Star spots are transient regions of cooler gas in the stellar surface and hence they appear darker than the surrounding surface regions. The formation of a large star spot on the stellar surface may cause a sufficient lowering in apparent brightness of the star as to be detectable, and this dip in apparent brightness will appear to come and go as the star rotates. However, a star spot is not a permanent feature and it will evolve over time eventually to disappear, resulting in the disappearance of the transit mimic. Although new

star spots may appear, they will most likely appear at different phases and stellar latitudes, resulting in transit mimics that come and go in a random fashion. The simplest way to identify such a transit mimic is to see if the transit signal disappears, changes in period, depth and/or duration over time. A good rule of thumb is that if the transit signal occurs at least 3 times with the same morphology, then it is unlikely to be due to star spots.

The next source of confusion comes from eclipsing binaries of various types. A typical detached eclipsing binary exhibits eclipses at a period of half of the true orbital period. A primary eclipse occurs when the cooler star (lower surface brightness) eclipses the hotter star (higher surface brightness) and a secondary eclipse occurs when the hotter star eclipses the cooler star. Since the stellar surface area obscured during a primary eclipse is the same as the area obscured during a secondary eclipse, the primary eclipse is the deeper of the two eclipses. For two main sequence stars, the hotter star is usually also the larger star. When the secondary object is a planet there is no secondary eclipse since the planet is a non-luminous object. Hence, the presence of eclipses of different depths in a lightcurve is a sure sign of an eclipsing stellar binary and photometric observations must cover the phases at half of the primary eclipse period in order to test this scenario.

This leads naturally into the case when the secondary eclipse is too shallow to be detected in the photometric data. This may occur when the amount of light eclipsed from the secondary star is much less than the amount of light eclipsed from the primary star (due to a combination of different sizes and surface brightnesses). Examples of such systems include a late M dwarf orbiting an F star or a Sun like star orbiting a giant star. There are various ways to detect these transit mimics. A moderate resolution spectrum of the entire visible wavelength range may be used to obtain the spectral type of the primary star, providing an estimate of its size. If the primary star is found to be too big (a giant for instance), then the companion may be too big to be a planet ( $R_p \geq 2R_J$ ) based on the value of  $R_p/R_*$  determined from the transit fit to the lightcurve. Dreizler et al. (2002) used this method to rule out the planetary status of a subset of the OGLE transit candidates.

If the radius of the host star is known, then one can determine the radius of the companion from the transit lightcurve. However, a companion with a radius similar to that of a hot Jupiter is not necessarily a planet. Hot Jupiters, brown dwarfs and late M dwarf

stars all have similar radii but very different masses (Perryman 2000). Therefore the best way to distinguish between these types of companions is to go directly to RV follow-up measurements. The drawback is that RV follow-up of transit candidates generally requires long exposure times on large telescopes in order to detect the radial velocity variations. A useful equation for the integration time per RV measurement  $t_{\text{obs}}$  in minutes required to achieve a precision of a quarter of the amplitude of the RV oscillations is supplied by Charbonneau (2003):

$$t_{\text{obs}} = 0.0363 \left( \frac{M_{\text{p}}}{M_{\text{J}}} \right)^{-2} \left( \frac{M_{*}}{M_{\odot}} \right)^{4/3} \left( \frac{P}{3.0 \text{ days}} \right)^{2/3} 10^{0.4(V-8)} \text{ min} \quad (2.36)$$

where  $V$  is the  $V$  band magnitude of the host star. This equation applies to the 10-m Keck telescope with the HIRES spectrograph. It is obvious from Equation 2.36 that the required integration time is very sensitive to planetary mass and stellar brightness. For example, a  $1.0M_{\text{J}}$  planet in a 3.0 day orbit around a Solar mass star with  $V = 10.0$ ,  $V = 14.0$  or  $V = 18.0$  requires an exposure time per measurement of  $\sim 0.23$  min,  $\sim 9.1$  min or  $\sim 6$  h respectively whereas a  $0.5M_{\text{J}}$  planet in a 3.0 day orbit around a Solar mass star with  $V = 10.0$ ,  $V = 14.0$  and  $V = 18.0$  requires an exposure time per measurement of  $\sim 0.92$  min,  $\sim 36$  min or  $\sim 25$  h. Clearly, radial velocity measurements can require a lot of telescope resources and accurate determination of the planetary mass may not even be possible for the fainter targets. With many transit candidates, the situation becomes unmanageable. It is therefore necessary to use other methods that are less resource intensive in order to rule out the transiting planet scenario for as many transit candidates as possible before resorting to RV follow-up.

Multi-colour photometry of the eclipse event will reveal if the eclipse has a waveband dependent eclipse depth. A planetary transit event is achromatic since the planet is non-luminous (although limb darkening effects may produce a very small colour change during the transit). However, a stellar binary eclipse is a chromatic event for any two stars with different colours. Hence, any detection of a waveband dependent eclipse depth indicates that the system is a stellar binary.

The presence of ellipsoidal variations or heating effects in the out of transit regions of the lightcurve indicate that the companion is massive or luminous respectively (Drake 2003).

Ellipsoidal variations are due to the host star being tidally distorted into an ellipsoidal shape by the companion and rotationally synchronised. A planetary companion is not massive enough to distort the shape of the star. The heating effect is due to heating on one side of the companion caused by irradiation by the host star. Again, a planet does not emit enough radiation to be detectable in the lightcurve. Therefore the presence of these specific variations in the observed lightcurve indicates that the companion is stellar.

If the secondary eclipse is of very similar depth (and shape) to the primary eclipse, which is generally the case for a grazing eclipsing stellar binary in which the two stars are very similar, then the differences between the primary and secondary eclipses may be undetectable in the photometric data. As a result, all of the eclipses may be mistaken as primary with a non-visible secondary eclipse. Neither spectral typing nor multi-waveband time-series photometry will help in this situation. However, the eclipses of such a stellar binary tend to be V shaped compared to the relatively flat bottomed annular eclipses. Careful modelling of the photometric data may therefore allow the transiting planet scenario to be ruled out.

Perhaps the most insidious transit mimic is that of a blend of a stellar binary with the light from a third star, not necessarily physically associated with the stellar binary itself. The eclipse depth of the stellar binary is diluted down by the light of the third star, leading to an underestimate of  $R_p/R_*$  when the lightcurve is fitted with the transiting planet model. There are a number of ways of testing this scenario including the use of adaptive optics to look for close neighbouring stars and careful analysis of the spectral line bisectors to look for small asymmetries in the spectral lines indicative of the presence of light from another star (Konacki et al. 2003a; Torres et al. 2004).

### 2.4.7 Conclusions

The probability of observing a planetary transit for a given star hosting an extra-solar planet is maximised for the class of planets known as hot Jupiters ( $P_{\text{tra}} \approx 10\%$ ). The transit depth is also maximised for hot Jupiters. Combining this with the fact that hot Jupiters have orbital periods of the order of a few days and transit durations that fit within



a single night of observations makes a transit survey most sensitive to this type of planet. Assuming that  $\sim 1\%$  of main sequence stars harbour a hot Jupiter (Butler et al. 2000), then a transit survey has the capacity to reveal  $\sim 1$  transiting hot Jupiter for every 1000 main sequence stars observed.

CCD imaging provides the means to photometrically monitor many stars at once. By carefully choosing the field to be observed and by using an instrument with a large field of view for the telescope (a CCD mosaic imager for example), one may observe thousands of stars at once. Photometric observations need to be of high cadence ( $\geq 8$  data points per hour) and high accuracy ( $\lesssim 1\%$  accuracy per data point) in order to maximise the signal to noise achieved. This is also important for defining the shape of any observed eclipses. Ideally more than 1 transit needs to be observed for a firm detection of a transit signal which in turn requires observing runs of  $\sim 10$ – $20$  nights for the detection of hot Jupiters in any particular field.

A transit survey will detect many transit candidates, the majority of which will be transit mimics involving eclipsing binary stars. Careful analysis of the lightcurves and simple follow-up observations will identify most of the mimics allowing the more time consuming radial velocity measurements to confirm or refute the planetary status of the (hopefully) few remaining transit candidates.

## 2.5 Summary

In this chapter we have given an overview of three of the most important techniques currently used to detect extra-solar planets: radial velocity, direct imaging/reflected light and gravitational microlensing. We have described the type of signal being looked for, how it is detected and what success the technique has had and/or is expected to have in the near future.

The transit technique has been reviewed with special attention since it is central to this thesis. We have considered in detail the theoretical aspects of a transit event and presented a program (*transitcurve.pro*) that we have developed in order to calculate theoretical transit lightcurves. We have also described the history of the technique and the difficulties that

---

are faced in proving the planetary nature of any one transit candidate. We have noted that transit surveys favour the detection of hot Jupiter type planets and we have described what is required of a transit survey in order to find these planets.

# 3

---

## A Transit Survey Of The Field Of Open Cluster NGC 7789

The University of St Andrews Planet Search (UStAPS) started to monitor open clusters in the search for planetary transits in 1999. Time series photometric data in a single waveband has been obtained on a total of three open clusters (NGC 6819, NGC 6940 and NGC 7789) during three observing runs. The first two observing runs (in 1999) were used to observe all three clusters in rotation, whereas the third run (in 2000) was used to observe NGC 7789 alone. This thesis deals with the NGC 7789 data from all three runs.

For details of the data reduction and results on NGC 6819 see Street et al. (2002) and Street et al. (2003). To summarise, Street et al. (2003) found 11 transit candidates but ruled out the planetary status of all of them except for one of the single-transit candidates. It was stated that they expected to detect  $\sim 11$  transiting hot Jupiters in the data if hot Jupiters are as common in this field as in the Solar neighbourhood, possibly indicating a lack of hot Jupiter type planets. Follow-up observations are yet to be carried out on any of the transit candidates. Street et al. (2002) report on the many new variable stars discovered by the survey. The data on NGC 6940 is still under analysis and will be published in the near future.



### 3.1 Introduction

The study of open clusters for transiting planets has a number of advantages over fields in other parts of the sky or Galactic plane. While providing a relatively large concentration of stars on the sky (but not so large as to cause blending problems as in the case of globular clusters observed from the ground), they also provide a set of common stellar parameters for the cluster members. These are metallicity, age, stellar crowding and radiation density. Also, the fainter cluster members are smaller stars and therefore they are likely to show deeper transit signatures, helping to offset sky noise contributions. The identification of the cluster main sequence in the colour magnitude diagram allows the assignment of a model-dependent mass and radius to each photometric cluster member, and assuming a law relating extinction to distance for the field allows the assignment of a model-dependent mass, radius and distance to all stars in the field under the assumption that they are main sequence stars. Transit candidates with well defined phased lightcurves may therefore be analysed in detail as to whether they are consistent with a transiting planet model. An estimate of the fraction of stars hosting a hot Jupiter (referred to as the hot Jupiter fraction) may be obtained by comparing the number of hot Jupiters that are actually detected to how many one would expect to detect using the knowledge of the star properties and the lightcurves themselves. The dependence of the hot Jupiter fraction on the cluster parameters may then be investigated by extending the experiment to other open clusters.

The observations of open cluster NGC 7789 and the data reduction process is the subject of this chapter. The main parameters of the cluster are shown in Table 3.1. For a good review of previous relevant work on this cluster see Gim et al. (1998).

### 3.2 Observations

We observed the open cluster NGC 7789 using the 2.5m Isaac Newton Telescope (INT) of the Observatorio del Roque de los Muchachos, La Palma, in the Canary Islands during three bright runs with dates 1999 June 22-30, 1999 July 22-31 and 2000 September 10-20. For brevity, these runs shall be referred to from now on as 1999-06, 1999-07 and 2000-09 respectively. We used the Wide Field Camera (WFC) which consists of a 4 EEV CCD

Table 3.1: Properties of the open cluster NGC 7789. Data taken from <http://obswww.unige.ch/webda> by Mermilliod, J.C. and the SIMBAD database.

RA (J2000.0)	23 <sup>h</sup> 57 <sup>m</sup>
Dec (J2000.0)	+56° 43'
$l$	115°48
$b$	-5°37
Distance (pc)	2337
Radius	~16'
Age (Gyr)	1.7
[Fe/H]	-0.08
$E(B - V)$	0.217

mosaic where each CCD is 2048x4096 pixels (Walton et al. 2001). The pixel scale is 0.33"/pix and field of view  $\sim 0.5^\circ \times 0.5^\circ$ . The gain and readout noise values for each chip were calculated automatically during the preprocessing stage of the data reduction (see Section 3.3.1). The mosaic field was centred on NGC 7789 at  $\alpha = 23^{\text{h}}57^{\text{m}}30^{\text{s}}$  and  $\delta = +56^\circ 43'41''$ .

The usual procedure for each night was to obtain  $\sim 5$  bias frames and  $\sim 8$  sky flat frames at both the beginning and end of the night. Observations on NGC 7789 in the runs 1999-06 and 1999-07 consisted of  $\sim 6$  pairs of 300s exposures taken every  $\sim 50$  minutes during the later part of each night. Observations in the 2000-09 run consisted of sequences of ten consecutive 300s exposures followed by a bias frame, repeated continuously throughout the whole of each night. With a readout time of 100s and various losses due to bad weather/seeing and telescope jumps, this resulted in a total of  $880 \times 300$  s exposures in Sloan  $r'$  over the three runs, with 691 of these exposures from the 2000-09 run alone. During the 2000-09 run, we also took 5 images of NGC 7789 with varying exposure times in Sloan  $i'$ , along with 5 sky flat frames, in order to provide us with the necessary colour information.

### 3.3 CCD Reductions

#### 3.3.1 Preprocessing: CCD Calibrations

Each run and each chip was treated independently for the purpose of the reductions. The reduction process was carried out by a single C-shell/IRAF script that runs according to a user-defined parameter file. Bad pixels were flagged in a user-defined detector bad pixel mask, and ignored where relevant. The script carries out the following steps:

1. Calculates the Heliocentric Julian Date (HJD) at mid exposure for each science frame.
2. Sorts the FITS files by observation type (bias, flat or science frame).
3. Sorts the flat frames and science frames by the filter used.
4. Deducts a  $3\sigma$ -clipped mean of the overscan region from each bias frame and then combines all the bias frames via a  $3\sigma$ -clipped mean to create a masterbias frame. The readout noise of the chip is determined from a  $3\sigma$ -clipped mean of the standard deviation image produced alongside the masterbias frame. The masterbias frame is then trimmed.
5. For each filter the following procedure is carried out on all flat frames. The script deducts a  $3\sigma$ -clipped mean of the overscan region from the flat frame, trims the frame, deducts the masterbias frame and calculates a  $3\sigma$ -clipped mean of a  $200 \times 200$  pixel square in the centre of the frame. The flat frames are then scaled to the same exposure level (using this mean value of the central  $200 \times 200$  pixels) so that they can be combined using the median pixel value. The resulting combined frame is a first approximation to the normalised master flat frame. Each original bias corrected flat frame is then divided by this initial normalised flat frame and a plane  $\Pi(x, y) = A + Bx + Cy$  is fitted to the resulting image. The plane  $\Pi(x, y)$  is now divided back into the original bias corrected flat frame in order to remove any non-uniform lighting effects (to first order) over the wide field and in order to normalise the frame. The corrected and normalised flat frames are then combined by the median pixel value to produce the normalised master flat frame  $M_F(x, y) \sim 1$ .

Table 3.2: Readout noise (ADU) and gain ( $e^-/\text{ADU}$ ) values for each chip.

Chip No.	Readout Noise (ADU)	Gain ( $e^-/\text{ADU}$ )
1	3.02	1.44
2	3.10	1.46
3	2.99	1.24
4	3.28	1.45

6. The gain is determined as follows from the flat frames for the Sloan  $r'$  filter. For a bias corrected flat frame  $F(x, y)$  and the corresponding plane  $\Pi(x, y)$  fitted in step 5, a model flat frame  $\bar{F}(x, y)$  is constructed such that  $\bar{F}(x, y) = \Pi(x, y) M_F(x, y)$ . The script then constructs the difference frame  $D(x, y)$  from  $D(x, y) = F(x, y) - \bar{F}(x, y)$ . Five  $100 \times 100$  pixel square regions along one of the diagonals of the difference frame are used to calculate values for the variance in  $F(x, y)$  which are paired up with the  $3\sigma$ -clipped means of the corresponding regions in the flat frame  $F(x, y)$ . After repeating this for all the flat frames, a plot of variance versus signal is made including the value of the read out noise squared for a signal of zero. A straight line is fitted to the data, optimally weighted using the formal uncertainty in the variance, by the method of  $\chi^2$  minimisation, the gradient being the inverse of the chip gain ( $e^-/\text{ADU}$ ).
7. For each filter the following procedure is carried out on all science frames. Saturated pixels are flagged in each science frame and added to the detector bad pixel mask to create a tailor made bad pixel mask for that frame. The script then deducts a  $3\sigma$  - clipped mean of the overscan region from the science frame, trims the frame, deducts the masterbias frame and divides the frame by the appropriate normalised master flat frame.

Table 3.2 shows the values of the readout noise and gain for each chip as calculated by the reduction script, and Figure 3.1 shows the data used to calculate the gain for each chip. Figure 3.2 shows a plot of some of the diagnostic data output by the reduction script for Chip 4.

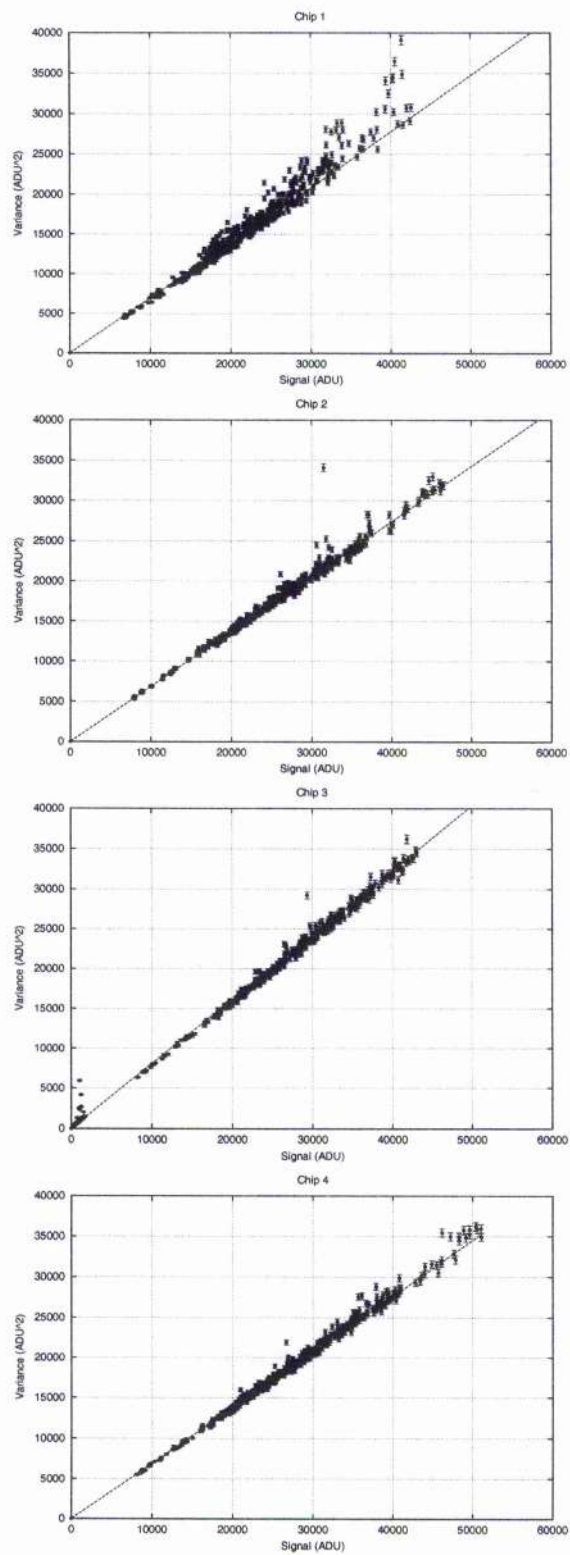


Figure 3.1: A plot of variance versus signal for each chip. The dashed line is the fit.



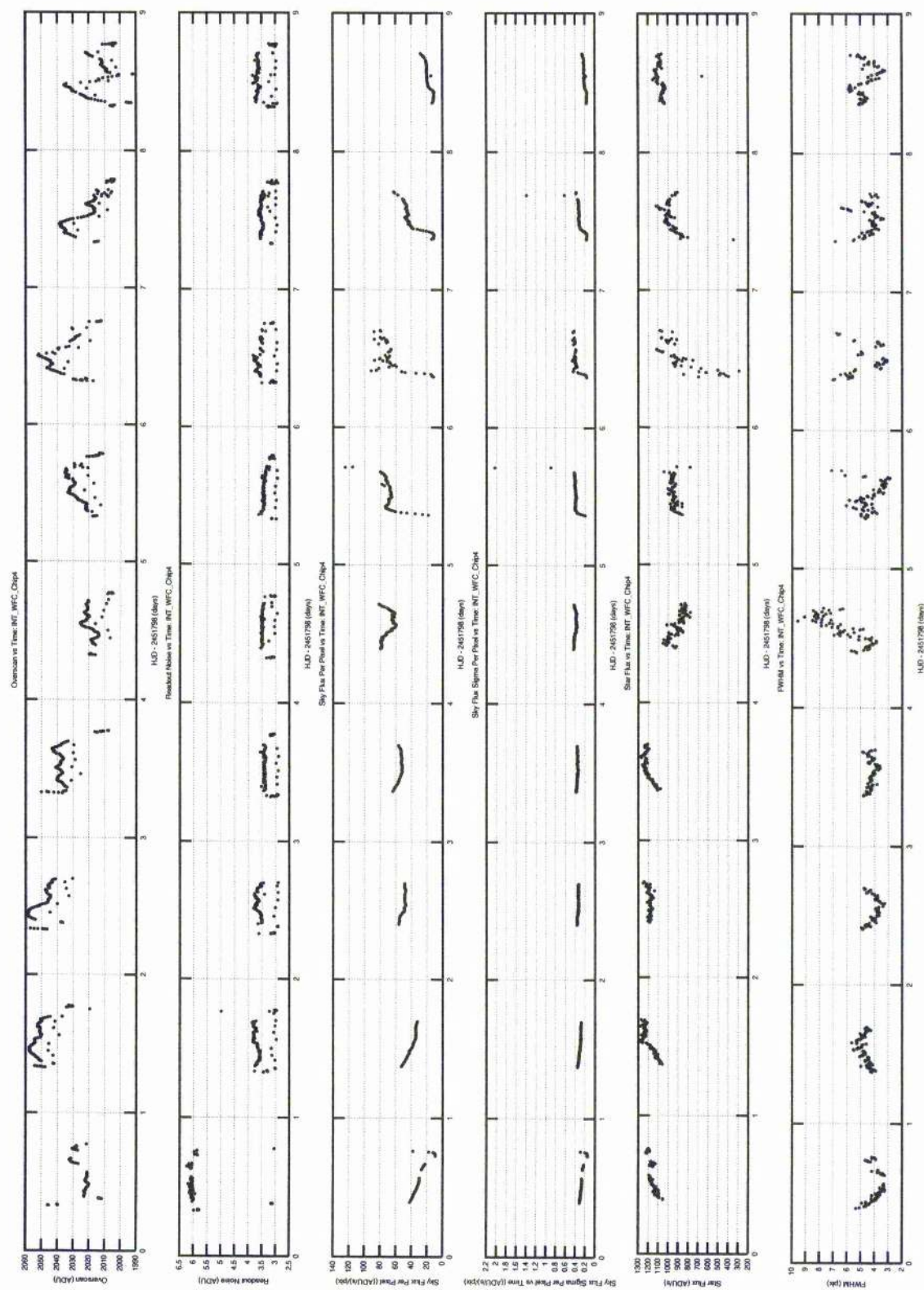


Figure 3.2: Diagnostic data for Chip 4.

### 3.3.2 Photometry: Difference Image Analysis

Differential photometry on the reduced science frames in the Sloan  $r'$  filter was accomplished using the method of difference image analysis (Alard & Lupton 1998; Alard 2000). Our implementation of this procedure was adapted from the code written for the MOA project (Bond et al. 2001), and it consists of three automated scripts. Bad pixels are ignored in the operations that the scripts perform.

1. The first script constructs a reference frame from selected frames with good seeing, and a star list from the reference frame. First, stars are detected and matched between the best seeing frames in order to derive a set of linear transformations and geometrically align the frames. The frames are then combined into a mean reference frame using the exposure times of the individual images as weights. The reference frame is analysed using IRAF's DAOPhot package (Stetson 1987). The package identifies stars on the reference frame and chooses a set of 175 point spread function (PSF) stars. A "penny2" PSF function that varies quadratically with position, along with a lookup table of residuals, is solved for. The neighbours of the PSF stars are then subtracted using this solution, and a new PSF function is solved for. This new solution is used to measure the instrumental fluxes and positions of all stars on the reference frame. The result is a reference frame with a corresponding star list. We used 13 consecutive best seeing images ( $\sim 1''$ ) to construct the reference frame.
2. The reference frame is used to produce a set of difference images. The main idea behind difference image analysis is that an image frame  $I(x, y)$  is related to the reference frame  $R(x, y)$  via the following equation:

$$I(x, y) = R \otimes K(x, y) + B(x, y) \quad (3.1)$$

where

$$R \otimes K(x, y) \equiv \int \int R(x - u, y - v) K(u, v, x, y) du dv \quad (3.2)$$

Here  $B(x, y)$  represents the change in the sky background, and  $K(u, v, x, y)$  is a spatially-varying convolution kernel relating the point-spread function on the reference frame to the point-spread function on the image frame at spatial position  $x, y$ .

We model the convolution kernel

$$K(u, v, x, y) = \sum_i a_i(x, y) b_i(u, v) \quad (3.3)$$

as the sum of a set of basis functions  $b_i(u, v)$  each formed as a product of a two-dimensional Gaussian function of  $u$  and  $v$  with a polynomial of degree 2 in  $u$  and  $v$ . For the basis functions we use 3 Gaussian components with sigmas of 2.1 pix, 1.3 pix and 0.7 pix and associated polynomial degrees of 2, 4 and 6 respectively. To allow for the kernel's spatial dependence, the coefficients  $a_i(x, y)$  are polynomials of degree 2 in  $x$  and  $y$ . The kernel is also normalised to a constant integral over  $u$  and  $v$  for each  $x$  and  $y$ , thus ensuring a constant photometric scale factor between the reference frame and image frame. We model the differential sky background  $B(x, y)$  as a polynomial of degree 2 in  $x$  and  $y$ . We solve for  $K(u, v, x, y)$  and  $B(x, y)$  in the least-squares sense for each science frame by fitting to pixel boxes around selected bright stars distributed uniformly across the reference frame. The kernel is assumed to be independent of  $x$  and  $y$  within each box. A difference image is then constructed for each science frame by rearranging Equation 3.1 to the following form and using the solutions for  $K(u, v, x, y)$  and  $B(x, y)$ :

$$D(x, y) = I(x, y) - R \otimes K(x, y) - B(x, y) \quad (3.4)$$

The difference image  $D(x, y)$  should simply be an image representative of the Poisson noise in  $I(x, y)$ . However, any objects that have varied in brightness in comparison to the reference frame should show up as positive or negative pixel areas on the difference image which may be measured to obtain the differential flux. In our analysis, each chip was split up into 8 square sections and the difference image constructed from solving for the kernel and differential sky background in each section. Also, a high signal-to-noise empirical PSF for the reference frame is constructed in each section by stacking up a set of stamps centred on suitable bright stars.

3. The third script measures the differential flux on each difference image via optimal PSF scaling at the position of each star. The normalised and sky-subtracted empirical PSF constructed for each square section of the reference frame in step 2 is convolved



with the kernel corresponding to the current difference image section. The convolved PSF is optimally scaled, at the position of the current star, to the difference image. A  $3\sigma$  clip on the residuals of the scaling is performed, and one pixel rejected. The scaling and rejection is repeated until no more pixels are rejected. The differential flux is measured as the integral of this scaled PSF.

A lightcurve for each star was constructed by the addition of the differential fluxes to the star fluxes as measured on the reference frame. The following equations were used:

$$f_{\text{tot}}(t) = f_{\text{ref}} + \frac{f_{\text{diff}}(t)}{p(t)} \quad (3.5)$$

$$m(t) = 25.0 - 2.5 \log(f_{\text{tot}}(t)) \quad (3.6)$$

where  $f_{\text{tot}}(t)$  is the star flux (ADU/s) at time  $t$ ,  $f_{\text{ref}}$  is the star flux (ADU/s) as measured on the reference frame,  $f_{\text{diff}}(t)$  is the differential flux (ADU/s) at time  $t$  as measured on the difference image,  $p(t)$  is the photometric scale factor (the integral of the kernel solution over  $u$  and  $v$ ) at time  $t$  and  $m(t)$  is the magnitude of the star at time  $t$ . Uncertainties are propagated in the correct analytical fashion.

Flux measurements were rejected for a  $\chi^2 \text{pix}^{-1} \geq 5.0$  for the PSF scaling, and for PSFs with a  $\text{FWHM} \geq 7.0$  pix, in order to remove bad measurements. Hence, all the stars have differing numbers of photometric measurements. In each run, lightcurves with less than half of the total possible epochs were rejected. For the 2000-09 run this analysis produced 8631 lightcurves on Chip 1, 7625 lightcurves on Chip 2, 8411 lightcurves on Chip 3 and 8830 lightcurves on Chip 4 (centred on the cluster). Figure 3.3 shows a diagram of the RMS scatter in the lightcurves against instrumental magnitude for the 2000-09 run for each chip. Similar diagrams were produced for the 1999-06 and 1999-07 runs but are not shown here for brevity.

Since each run was treated independently for the reductions, each chip has three different reference frames and hence each star has three different reference magnitudes. For a particular star, let us denote the reference magnitude from the 2000-09 run minus the reference magnitude from the 1999-06 run by  $\Delta m_1$  and the reference magnitude from the

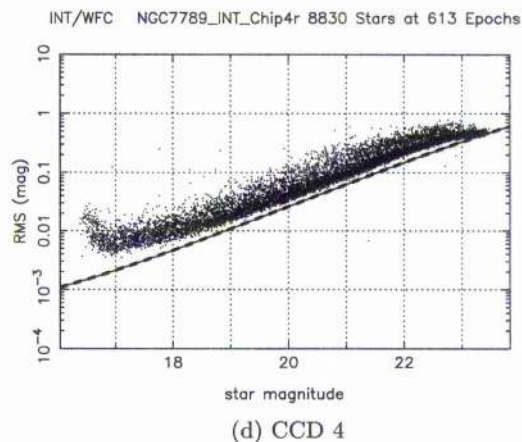
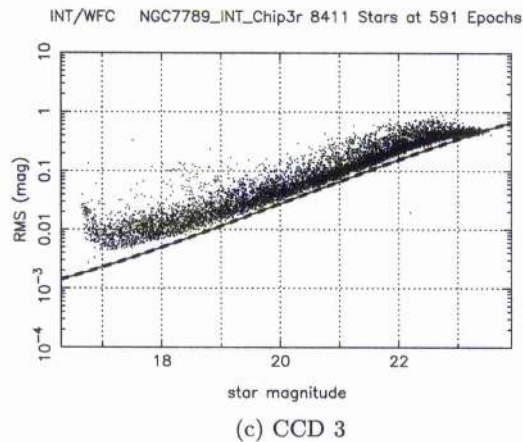
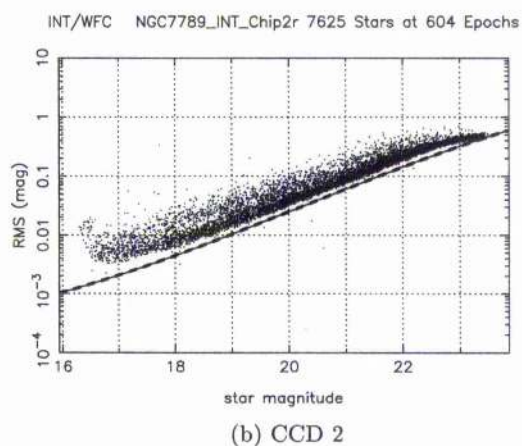
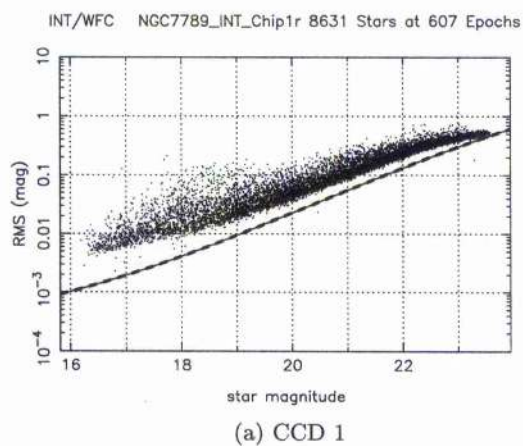


Figure 3.3: Plots of standard deviation (RMS) of the lightcurves against mean instrumental Sloan  $r'$  magnitude for all stars from each chip for the 2000-09 run. The lower curve in each diagram represents the theoretical noise limit for photon and readout noise.

2000-09 run minus the reference magnitude from the 1999-07 run by  $\Delta m_2$ . For each chip, we have calculated the unweighted mean of  $\Delta m_1$  and  $\Delta m_2$  over all stars on that chip. We then added the resulting  $\overline{\Delta m_1}$  and  $\overline{\Delta m_2}$  to the lightcurve data points in the 1999-06 and 1999-07 runs respectively. The values of the means  $\overline{\Delta m_1}$  and  $\overline{\Delta m_2}$  for each chip along with the standard deviations about the means  $\sigma_1$  and  $\sigma_2$  respectively are presented in Table 3.3.

As can be seen from Figure 3.3, we have obtained high precision photometry with an RMS accuracy of  $\sim 3\text{--}5\text{mmag}$  at the bright end. Most stars are limited by sky noise because all three runs were during bright time. However, the “backbone” of points on each diagram lies above the theoretical limit by a factor of  $\sim 1.5\text{--}2.0$  depending on the chip being considered. We put this down to systematic errors in the data due to a subset of low quality difference images and/or sections of difference images that were produced from science frames taken on nights of poor quality seeing/atmospheric conditions.

### 3.4 Astrometry And Colour Data

#### 3.4.1 Astrometry

Astrometry was undertaken by matching 358 stars from the four reference frames (one for each chip) with the USNO-B1.0 star catalogue (Monet et al. 2003) using a field overlay in the image display tool GAIA (Draper 2000). The WFC suffers from pincushion distortion, hence it was necessary to fit a 9 parameter astrometric solution to the reference frames in order to obtain sufficiently accurate celestial coordinates for all the stars. The 9 parameters are made up of 6 parameters to define the linear transformation between pixel coordinates and celestial coordinates, 2 parameters to define the plate centre and 1 parameter to define the radial distortion coefficient. The starlink package ASTROM (Wallace 1998) was used to do the fit and the achieved accuracy was  $\sim 0.4$  arcsec RMS radially for the 358 matching stars. The astrometric fit was then used to calculate the J2000.0 celestial coordinates for all stars with a lightcurve.

### 3.4.2 Colour Indices

The best image in the Sloan  $i'$  filter was aligned with the Sloan  $r'$  reference frame for each chip and the magnitudes of the stars were measured using DAOPhot PSF fitting in the same way as they were measured on the reference frame in Section 3.3.2. Table 3.4 shows the number of stars with lightcurves that have Sloan  $r' - i'$  colour indices as a result.

### 3.4.3 Colour Magnitude Diagrams

Figures 3.4(a)-3.4(d) show an instrumental colour magnitude diagram (CMD) for each chip. The cluster main sequence is clearly visible. Chip 4 is centred on the cluster and as expected shows the strongest cluster main sequence. A theoretical cluster main sequence is plotted on each diagram over the cluster main sequence. We have used the theoretical models of Baraffe et al. (1998) for the stellar mass range  $0.60M_{\odot} \leq M_* \leq 1.40M_{\odot}$ , the age of the cluster (1.7Gyr) and solar-type metallicity  $[M/H] = 0$  in order to predict the main sequence absolute magnitudes, colours and radii. Below a mass of  $0.60M_{\odot}$  the Baraffe model predicts  $R-I$  colours substantially bluer than the observed cluster main sequence, a limitation noted in Baraffe et al. (1998). As a result we used data from Lang (1992) for the stellar mass range  $0.08M_{\odot} \leq M_* \leq 0.60M_{\odot}$ . The combined model for the cluster main sequence supplies an absolute magnitude  $M_R$ , an absolute magnitude  $M_I$  and a stellar radius  $R_*$  for the stellar mass range  $0.08M_{\odot} \leq M_* \leq 1.40M_{\odot}$ . We interpolated this combined model with cubic splines.

The interstellar medium (ISM) in the Milky Way is mostly concentrated in the Galactic plane and the density law governing its mean distribution (ignoring small scale variations) can be modelled by an Einasto law:

$$\rho(R, z) = \rho_0 \exp \left( - \left( \frac{R - R_{\odot}}{h_R} \right) \right) \exp \left( - \frac{|z|}{h_z} \right) \quad (3.7)$$

where  $R$  is the Galactocentric distance,  $z$  is the height above the Galactic plane,  $\rho_0$  is the local density of the ISM,  $R_{\odot}$  is the distance of the Sun from the Galactic centre,  $h_R$  is the ISM density scale height in the  $R$  direction and  $h_z$  is the ISM density scale height in the  $z$  direction. One may derive the density  $\rho$  of the ISM as a function of distance  $d$  from the



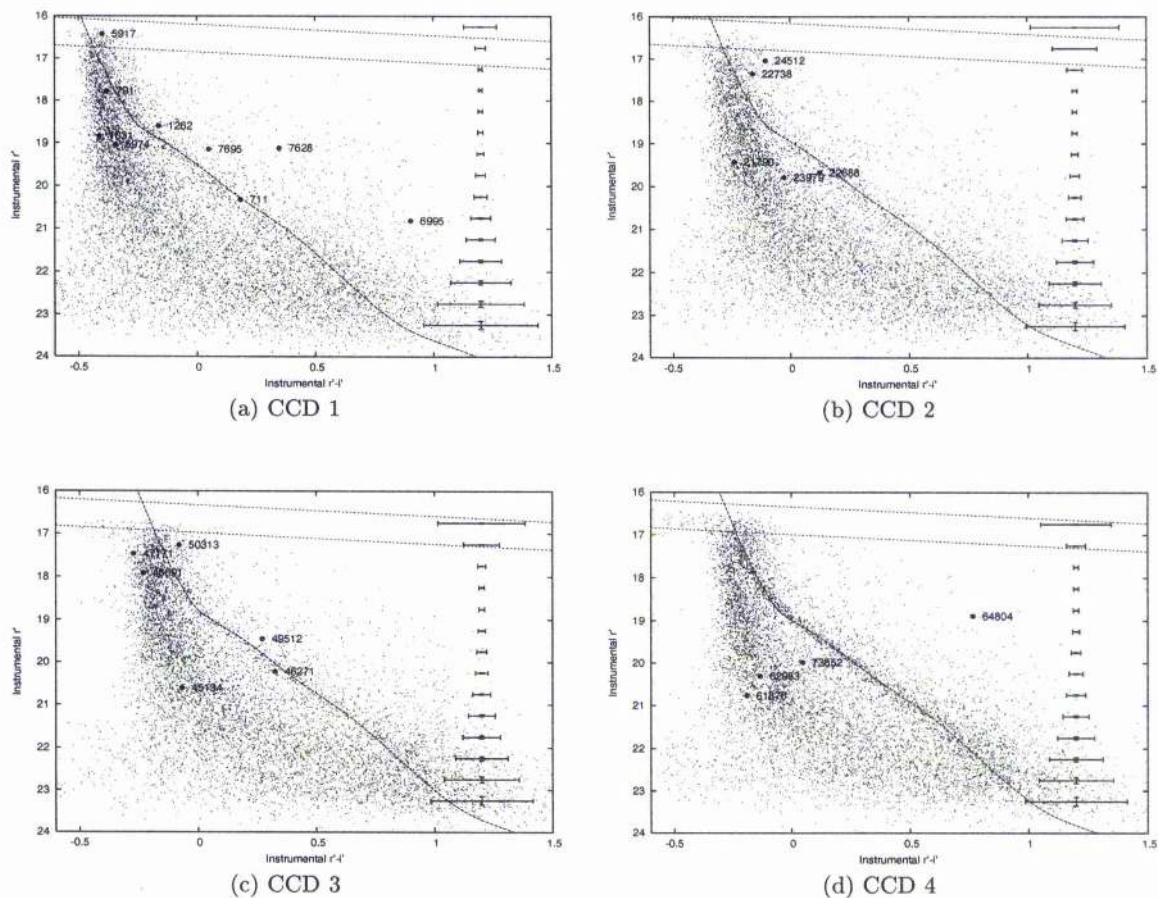


Figure 3.4: Instrumental CMDs for all stars from each chip for the 2000-09 run. The main sequence is visible on each chip, and the theoretical cluster main sequence is overlayed as the dashed line. The straight dotted lines are the faint limits for giant stars assuming no extinction and the law relating extinction to distance in Equation 3.8. The transit candidates of Section 4.2 are marked on as solid circles. The errorbars on the right hand side of each diagram represent the mean error bar on each measurement for 0.5 magnitude bins.

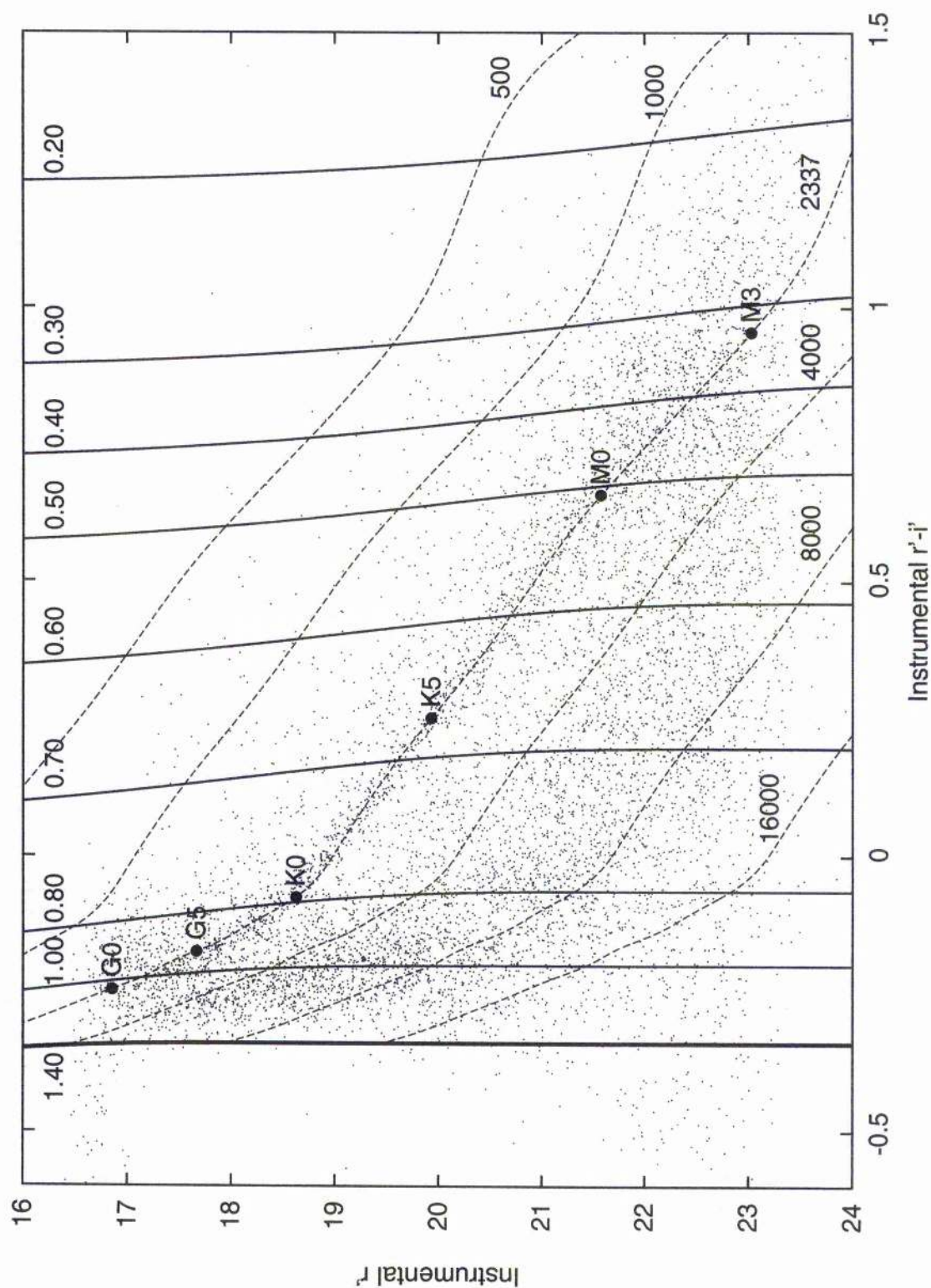


Figure 3.5: Instrumental CMD for Chip 4 (see the text in Section 3.4.4). The numbers along the top are masses in units of  $M_{\odot}$  and the numbers along the bottom and right are distances in parsecs.

Sun in the direction of the open cluster NGC 7789 by using trigonometrical arguments to rewrite  $R$  and  $z$  as functions of  $d$ . In this derivation, we have assumed that the Sun has Galactic coordinates  $(R, z) = (8.5\text{kpc}, 0.015\text{kpc})$  and that  $\rho_0 = 0.021 M_\odot \text{pc}^{-3}$ ,  $h_R = 4.5\text{kpc}$ ,  $h_z = 0.14\text{kpc}$  as given in Robin et al. (2003). In any wave band, the total extinction  $A$  as a function of  $d$  is proportional to the integral of  $\rho(d)$  over  $d$ . Hence, absorbing the constant  $\rho_0$  into a new constant  $K$  we have:

$$A(d) = K \int_0^d \rho(u) du \quad (3.8)$$

Adopting  $E(B - V) = 0.217$  for the cluster (Table 3.1), we calculate the corresponding extinction to be  $A_R \approx 0.547$  and  $A_I \approx 0.429$  in the  $R$  and  $I$  bands respectively, evaluated with a synthetic photometry code (XCAL) using a Galactic extinction curve from Seaton (1979). This extinction applies to stars at the cluster distance  $d_c = 2337\text{pc}$ , and hence, by numerically evaluating the integral in Equation 3.8, we may calculate values for  $K$  that apply to the  $R$  and  $I$  bands as  $K_R = 2.20 \times 10^{-2} \text{mag} M_\odot^{-1} \text{pc}^2$  and  $K_I = 1.73 \times 10^{-2} \text{mag} M_\odot^{-1} \text{pc}^2$  respectively.

We have used the law relating extinction to distance as given in Equation 3.8 to correct the absolute magnitudes  $M_R$  and  $M_I$  of the theoretical main sequence to the observed magnitudes  $R(d)$  and  $I(d)$  respectively. In the following equations, the distance  $d$  has units of parsecs (pc):

$$R(d) = M_R + 5 \log(d) - 5 + A_R(d) \quad (3.9)$$

$$I(d) = M_I + 5 \log(d) - 5 + A_I(d) \quad (3.10)$$

where  $A_R(d)$  and  $A_I(d)$  are versions of Equation 3.8 with  $K = K_R$  and  $K = K_I$  respectively.

Conversions between the Johnson-Cousins  $R$  and  $I$  magnitudes and the Sloan  $r'$  and  $i'$  magnitudes were done using the following predetermined relations presented on the Cambridge Astronomical Survey Unit (CASU) webpage<sup>1</sup>:

$$r' = R + 0.275(R - I) + 0.008 \quad (3.11)$$

$$r' - i' = 1.052(R - I) + 0.004 \quad (3.12)$$

<sup>1</sup><http://www.ast.cam.ac.uk/~wfcsur/index.php>



Due to the lack of observations of standard stars, it was necessary to fit the interpolated theoretical main sequence to the cluster main sequence on the CMD for each chip by eye, after correcting for the cluster distance and extinction, via simple  $r'$  and  $r' - i'$  offsets. These offsets are displayed in Table 3.5. Note that the required horizontal and vertical shifts are correlated, since shifts parallel to the main sequence would have no effect if the main sequence were a straight line. Fortunately, the kink (change of slope) in the main sequence near the spectral type K0 ( $0.8M_{\odot}$ ) allows us to estimate both vertical and horizontal shifts. This feature is clearly visible on all 4 chips.

#### 3.4.4 Stellar Masses, Radii And Distances

The identification of the cluster main sequence on each CMD allows a model-dependent mass, radius and distance for each star to be determined using the theoretical main sequence, assuming that each star is a main sequence star. Giant stars (MK Luminosity Class III) have absolute magnitudes in the range  $1.7 \leq M_V \sim M_R \leq -6.5$  (Lang 1992). Assuming that the Sun lies in the Galactic plane at a distance of 8.5kpc from the Galactic centre (the IAU value) and assuming that the Galactic disk has a radius of 14.0kpc (Robin, Cr   , & Mohan 1992), then the distance to the edge of the Galaxy in the direction of NGC 7789 may be calculated as  $\sim 8.1$ kpc using elementary trigonometry. The magnitude of the dimmest giant at 8.1kpc assuming no extinction is  $R = 16.2$  and, assuming the law relating extinction to distance in Equation 3.8, the dimmest giant has a magnitude of  $R = 16.9$ . These faint limits are marked on the CMDs in Figures 3.4(a)-3.4(d) as dotted lines. From this simple argument it can be seen that only the brightest stars in our sample will be contaminated with giant stars.

In principle, for each star on the CMD, it is possible to choose a value for the distance parameter  $d$  in Equations 3.9 and 3.10 such that the theoretical main sequence passes through the star's position on the CMD. The solution  $d = d_*$  is then the distance to the star. The star mass  $M_*$  and radius  $R_*$  may subsequently be determined from where the star lies on the theoretical main sequence at a distance of  $d_*$ . Figure 3.5 shows the grid of star masses and distances used for Chip 4. The solid vertical lines represent lines of constant



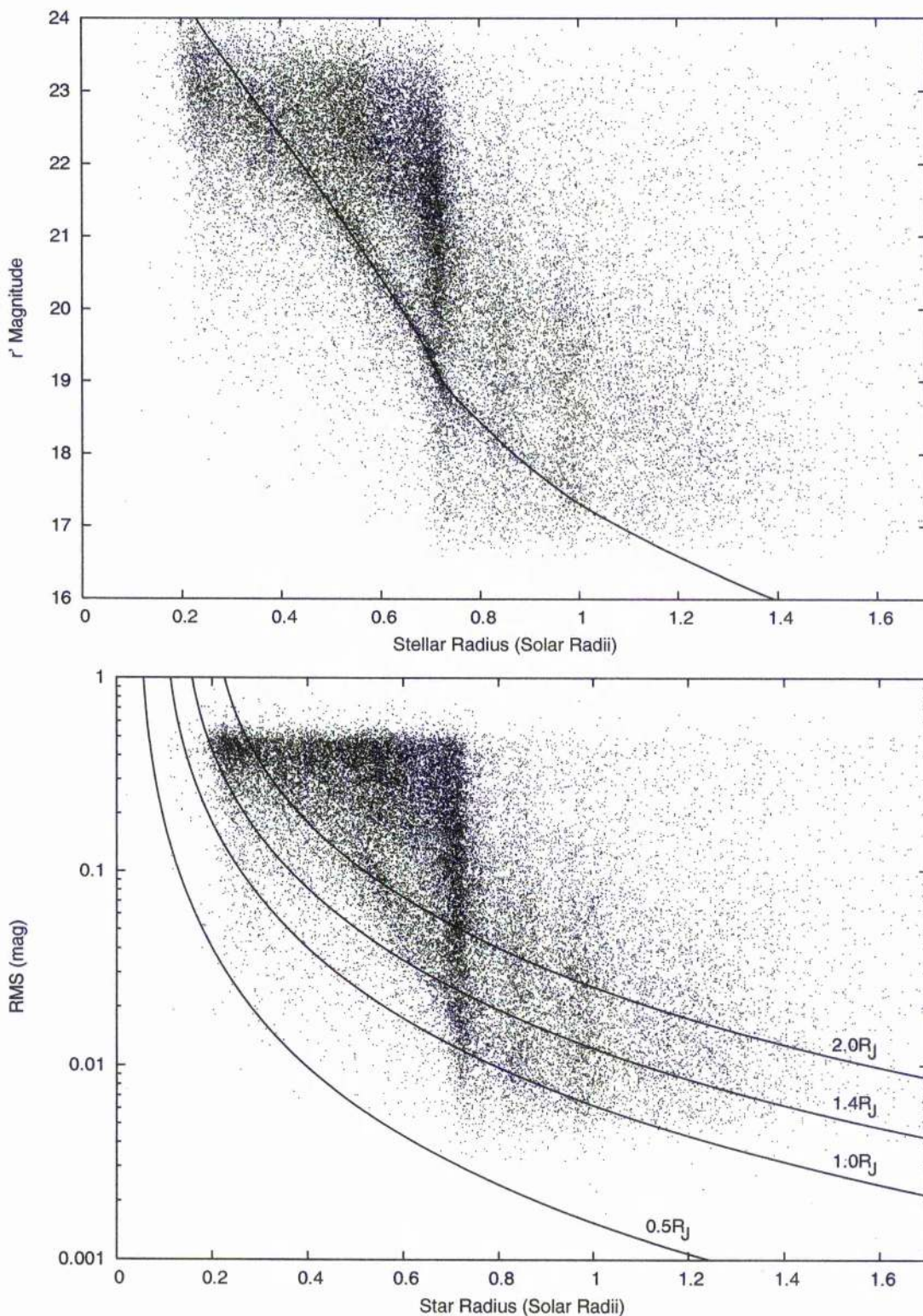


Figure 3.6: **Top:** A plot of  $r'$  magnitude against star radius for all four chips. The continuous line is the theoretical cluster main sequence. **Bottom:** Standard deviation (RMS) of the lightcurves against star radius for all four chips. The continuous curves are the detection limits for a planet of the quoted radius as a function of star radius (see Section 3.4.5).

Table 3.3: Magnitude offsets  $\overline{\Delta m_1}$  and  $\overline{\Delta m_2}$  added to the lightcurve data points from the 1999-06 and 1999-07 runs respectively.

Chip No.	$\overline{\Delta m_1}$	$\sigma_1$	$\overline{\Delta m_2}$	$\sigma_2$
1	-0.832	0.037	-0.753	0.041
2	-0.698	0.036	-0.843	0.043
3	-0.596	0.031	-0.404	0.054
4	-0.527	0.036	-0.470	0.037
Run:	1999-06		1999-07	

Table 3.4: Number of stars with a Sloan  $r'$  lightcurve, and the number of such stars with a Sloan  $r' - i'$  colour index.

Chip No.	No. Stars With A Lightcurve	No. Stars With A Lightcurve And An $r' - i'$ Colour Index	Percentage
1	8631	8497	98.4%
2	7625	7576	99.4%
3	8411	8290	98.6%
4	8830	8672	98.2%
Total:	33497	33035	98.6%

Table 3.5: Offsets determined by eye between observed  $r'_{\text{obs}}$  and  $r'_{\text{obs}} - i'_{\text{obs}}$  magnitudes, and theoretical  $r'(d_c)$  and  $r'(d_c) - i'(d_c)$  main sequence magnitudes for the cluster distance  $d_c$ .

Chip No.	$r'_{\text{obs}} - r'(d_c)$	$(r'_{\text{obs}} - i'_{\text{obs}}) - (r'(d_c) - i'(d_c))$
1	-0.3	-0.95
2	-0.3	-0.80
3	-0.1	-0.70
4	-0.1	-0.75
Estimated Error:	$\pm 0.1$	$\pm 0.05$

stellar mass (and radius), and are labelled at the top of the diagram in units of  $M_{\odot}$ . The “diagonal” dashed lines represent theoretical main sequence models at different distances, and the distances are labelled to the right and bottom of the diagram in units of parsecs. Fiducial spectral types are marked on the cluster theoretical main sequence for clarity.

Due to the steepness of the theoretical main sequence in the CMD for star masses greater than  $0.80M_{\odot}$ , the determined star properties become more uncertain above  $0.80M_{\odot}$ . Also, the theoretical main sequence that we have used terminates at a mass of  $1.40M_{\odot}$ , which leads to a small region where there are no solutions for  $d_*$ . In Figure 3.5, this region is blueward of the thick continuous line (corresponding to a mass of  $1.40M_{\odot}$ ). Stars with no solution for  $d_*$  have masses greater than  $1.40M_{\odot}$  (and radii greater than  $1.70R_{\odot}$ ) and large distances. It is around these stars that it is hardest to detect a transiting planet and hence a lack of solution for  $d_*$ ,  $M_*$  and  $R_*$  will hardly affect the completeness of our survey. Table 4.2 shows the star masses, radii and distances obtained by the above procedure for the transit candidates discussed in Section 4.2. The star  $r' - i'$  colours have been corrected where necessary for any lightcurve variations (since the reference frame from which the  $r'$  magnitude was determined has a different epoch to the  $i'$  frame from which the  $i'$  magnitude was determined).

In Figure 3.6 we plot the  $r'$  magnitude and standard deviation (RMS) of the lightcurve of each star versus the stellar radius derived from the main-sequence model and the observed  $r' - i'$  colour index. A vertical stripe of stars is evident with  $R_* \sim 0.75 R_{\odot}$ . This arises because of a relatively rapid change in the colour index with mass for the theoretical main sequence in this mass range. This effect is also evident in Figure 3.5, where the vertical iso-mass lines are more widely spaced for  $0.5M_{\odot} < M_* < 0.8M_{\odot}$ . If the mass function and mass-radius relationship for main sequence stars are both smooth, then this effect represents a deficiency in the  $R - I$  colour index of the stellar models. The mass-radius relationship for our theoretical main sequence is shown in Figure 3.7.

In order to test the theoretical main sequence that we have adopted, we compared it to observations of solar neighbourhood main sequence stars in the  $M_R$  versus  $R - I$  domain. The data on the main sequence stars were taken from Bessell (1990)<sup>2</sup>. The result is shown

<sup>2</sup>This data may be found at: <http://www-int.stsci.edu/~inr/cmd.html>

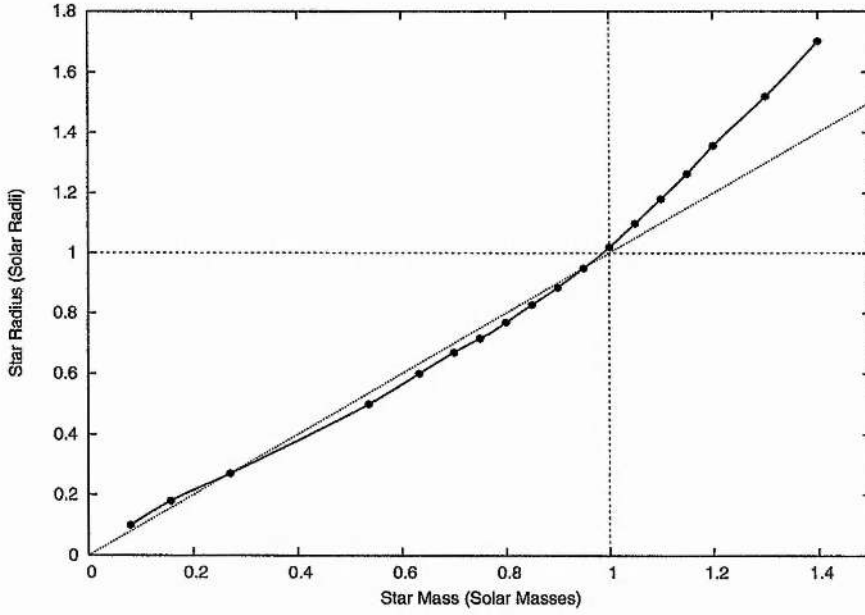


Figure 3.7: Mass-radius relationship for the theoretical main sequence (filled circles) along with the cubic spline (continuous line).

in Figure 3.8 where we plot our adopted theoretical main sequence (filled circles), the cubic spline (continuous line), the Baraffe et al. (1998) model (stars) and the observed main sequence stars (small filled squares). The numbers marked on the diagram are the star masses (in units of Solar mass) corresponding to our theoretical main sequence model. The diagram clearly shows the shortcomings of the Baraffe model for  $M_* < 0.60M_{\odot}$ , and we note that although there seems to be a good agreement between our adopted theoretical main sequence and the observations, there is a small discrepancy of  $\sim 0.02$  mag in  $R - I$ .

### 3.4.5 Number Of Expected Transiting Planets

We would like to estimate, for a specific signal-to-noise ratio, the accuracy required per photometric measurement in order to detect a transiting planet of radius  $R_p$  orbiting a star of radius  $R_*$ . By working in units of magnitude rather than units of flux, and by following the same reasoning as in Section 2.4.4, we can derive:

$$S/N = \frac{\Delta m}{\sigma} \sqrt{\frac{N_{\text{out}} N_{\text{in}}}{N_{\text{out}} + N_{\text{in}}}} \quad (3.13)$$



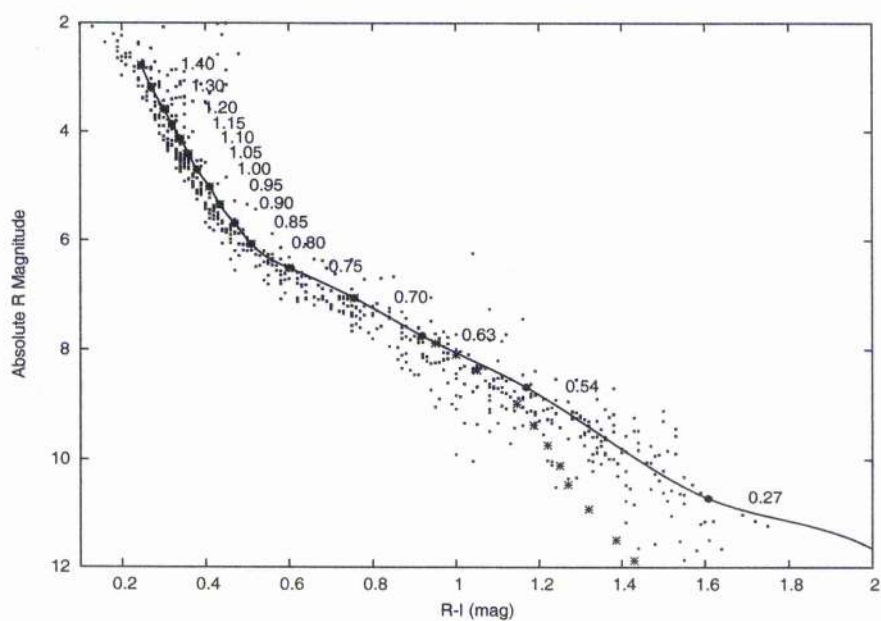


Figure 3.8: Plot of the theoretical main sequence (filled circles), the cubic spline (continuous line), the Baraffe et al. (1998) model (stars) and the observed main sequence stars (small filled squares) in the  $M_R$  versus  $R - I$  domain. The numbers marked on the diagram are the star masses (in units of Solar mass) corresponding to the theoretical main sequence model.

where  $\Delta m$  is the transit depth in magnitudes and  $\sigma$  is the uncertainty in each photometric measurement in magnitudes. The numbers  $N_{\text{out}}$  and  $N_{\text{in}}$  are the number of data points out-of-transit and in-transit respectively. Let us consider the limit when  $N_{\text{out}} \gg N_{\text{in}}$ . Then, from Equation 3.13, we have:

$$S/N \approx \frac{\Delta m \sqrt{N_{\text{in}}}}{\sigma} \quad (3.14)$$

Let  $m_0$  be the magnitude of the star out-of-transit and  $m_{\text{in}}$  be the magnitude of the star during transit, and let  $f_0$  be the star flux out-of-transit and  $f_{\text{in}}$  be the star flux during transit. Also, let  $z$  be the magnitude zero point. Then:

$$\begin{aligned} \Delta m &= m_{\text{in}} - m_0 \\ &= z - 2.5 \log(f_{\text{in}}) - z + 2.5 \log(f_0) \\ &= -2.5 \log\left(\frac{f_{\text{in}}}{f_0}\right) \\ &= -2.5 \log\left(1 - \frac{\Delta f}{f_0}\right) \end{aligned} \quad (3.15)$$

In the last line we have used Equations 2.28 and 2.29. Substitute Equation 2.26 into Equation 3.15 to get:

$$\Delta m = -2.5 \log\left(1 - \left(\frac{R_p}{R_*}\right)^2\right) \quad (3.16)$$

Finally, substitute Equation 3.16 into Equation 3.14:

$$S/N \approx \frac{-2.5 \sqrt{N_{\text{in}}}}{\sigma} \log\left(1 - \left(\frac{R_p}{R_*}\right)^2\right) \quad (3.17)$$

For random sampling of the orbital period  $P$ , the probability that a given data point catches a transit is  $\Delta t/P$ , where  $\Delta t$  is the transit duration. For a HD 209458b-like system ( $\Delta m \approx 15\text{mmag}$ ,  $\Delta t \approx 3\text{ h}$ ,  $P \approx 3.5\text{ d}$ ), this fraction is  $\Delta t/P = 3.5\%$  and thus  $N \sim 30$  of our 880 lightcurve data points would catch a transit. For our survey, a star with  $\sigma \approx 10\text{mmag}$  would allow detection of HD 209458b-like transits with  $S/N \approx 8$ . In Section 4.1 we adopt a more conservative transit detection threshold  $S/N \approx 10$ . Taking  $S/N = 10$  and  $N = 30$ , and given a value for  $R_p$ , we may calculate the required RMS accuracy  $\sigma$  as a function of stellar

radius  $R_*$  by rearranging Equation 3.17. These curves are plotted in Figure 3.6 (bottom) for planetary radii of  $0.5R_J$ ,  $1.0R_J$ ,  $1.4R_J$  and  $2.0R_J$ .

By counting the number of stars that lie beneath each of the continuous curves in Figure 3.6 (bottom), we obtain estimates for the number of stars that achieve the required accuracy in their lightcurves to enable detection of a planet of the quoted radius. Using this criterion, we expect our sample to contain 21 stars for which we can detect a planet of radius  $0.5R_J$ , 1451 stars for which we can detect a planet of radius  $1.0R_J$ , 5068 stars for which we can detect a planet of radius  $1.4R_J$  and 12080 stars for which we can detect a planet of radius  $2.0R_J$ . Assuming that  $\sim 1\%$  of main sequence stars have a  $1.4R_J$  hot Jupiter companion, and that  $\sim 10\%$  of such systems exhibit transits, then we may expect  $\sim 5$  stars in our sample to reveal planetary transits. In Chapter 5, we will model in more detail the number of hot Jupiters that we expect to detect from our survey using Monte Carlo simulations.

### 3.5 Summary

This chapter has been used to introduce the University of St Andrews Planet Search in open clusters and the motivation behind such a survey. More specifically we have reported on the observations of NGC 7789 that were made with the Isaac Newton Telescope in La Palma during three runs in 1999 and 2000. We have described the reduction procedures and photometry methodology that were employed on the data, which were subsequently developed into an efficient data pipeline for future use on similar data sets. The pipeline is split into a preprocessing C-shell/IRAF script that carries out the CCD calibrations, and a DIA package that carries out differential photometry.

We have described how we achieved the astrometry calibrations and obtained colour measurements for the stars. We have shown how we used the colour magnitude diagrams along with a theoretical main sequence model and a law relating extinction to distance in order to derive stellar masses, radii and distances for the majority of stars in our data set. Finally, we have estimated how many transiting hot Jupiter planets we expect to detect from the data by considering a simple signal-to-noise calculation.

## 21 Eclipsing Binaries And 3 Planetary Transit Candidates

At this stage we have  $\sim 33000$  stars with a lightcurve, and a model-dependent mass, radius and distance. Our next job in the search for hot Jupiters in the field of NGC 7789 is to detect any eclipses that might be present in the lightcurve data and to analyse these transit candidates in detail in order to determine whether they are consistent with a transiting planet model. Many of these transit candidates are actually eclipsing stellar binaries, although this is not obvious without carefully modelling the photometric data.

### 4.1 Transit Detection

We used a matched filter algorithm to search for transits in the lightcurves. Adopting a square “boxcar” shape for the transit lightcurve, the transit model has 4 parameters: the out-of-eclipse magnitude  $m_0$ , the time of mid-transit  $t_0$ , duration  $\Delta t$  and depth  $\Delta m$ . We search for transits with durations ranging from 0.5 h to 5 h, spanning this range with 12 values of  $\Delta t$  spaced by factors of 1.23. We move the transit centroid  $t_0$  through the data in steps of  $\Delta t/4$ . As illustrated in Figure 4.1, we fit both a constant and a boxcar transit lightcurve to the data points in a window of width  $5\Delta t$  centred on each value of  $t_0$ . Our



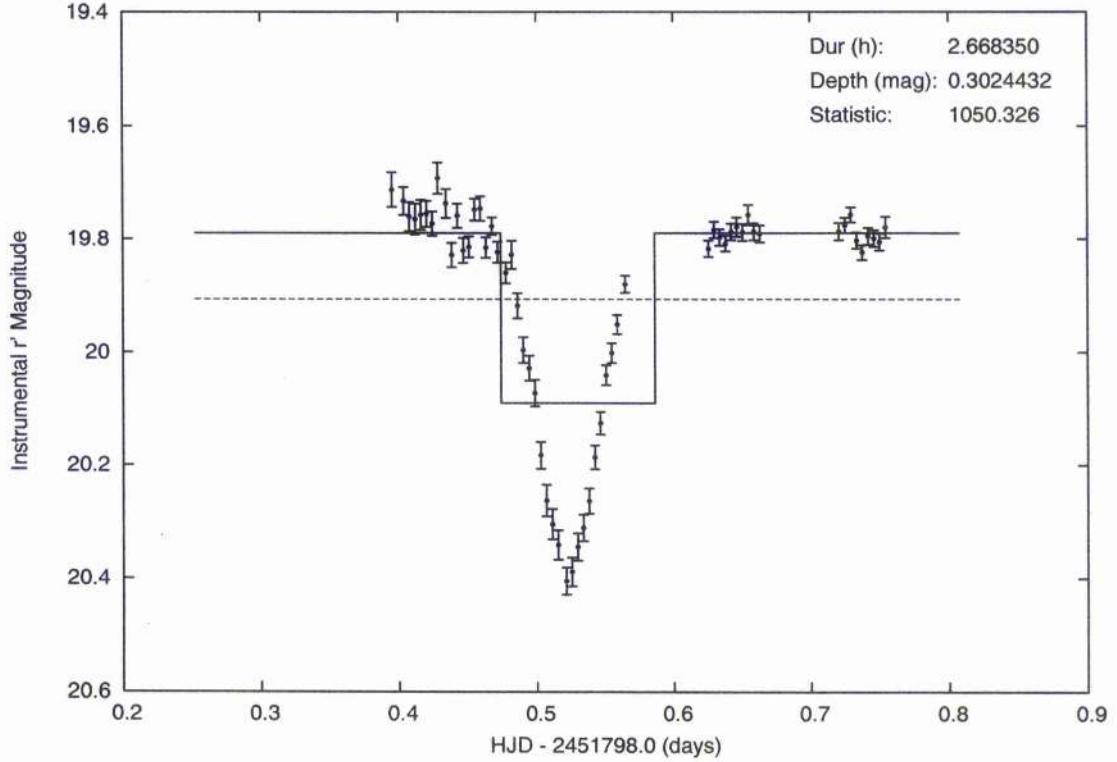


Figure 4.1: An example boxcar transit fit showing the in-transit and out-of-transit zones (continuous line) and the constant fit (dashed line). The horizontal axis is time (days) and the vertical axis instrumental Sloan  $r'$  magnitude. "Statistic" is the value of the transit statistic  $S_{\text{tra}}^2$  for this fit.

transit detection statistic is:

$$S_{\text{tra}}^2 \equiv \frac{\chi_{\text{const}}^2 - \chi_{\text{tra}}^2}{\left( \frac{\chi_{\text{out}}^2}{N_{\text{out}} - 1} \right)} \quad (4.1)$$

where  $\chi_{\text{tra}}^2$  is the chi squared of the boxcar transit fit,  $\chi_{\text{const}}^2$  is the chi squared of the constant fit,  $\chi_{\text{out}}^2$  is the chi squared of the boxcar transit fit for the  $N_{\text{out}}$  out-of-transit data points. The statistic  $S_{\text{tra}}^2$  is effectively the squared signal-to-noise ratio of the fitted transit signal renormalised to the reduced chi squared of the out-of-transit data points. This modified matched filter algorithm was designed to help downweight systematic errors with  $\chi_{\text{out}}^2 / (N_{\text{out}} - 1) > 1$  (and serendipitously, variables), since transit signals should have  $\chi_{\text{out}}^2 / (N_{\text{out}} - 1) \sim 1$ .

The transit detection algorithm outlined above was applied to the 1999-07 and 2000-09 runs. Initial tests with  $S_{\text{tra}}^2$  generated many spurious transit candidates in which the

Table 4.1: The number of raw transit candidates and remaining transit candidate lightcurves after the weeding process for each chip over the two runs.

Chip No.	No. Raw Transit Candidates	No. Remaining Transit Candidate Lightcurves
1	617	9
2	311	5
3	721	6
4	533	4
Total:	2182	24

transit fit matched low data points at the beginning or end of a night. To suppress these we introduced additional requirements on the number of in-transit and out-of-transit lightcurve data points. For the densely sampled 2000-09 run, we required at least 3 in-transit and 8 out-of-transit lightcurve data points for a transit detection. For the more sparsely sampled 1999-07 run, we required at least 2 in-transit and 6 out-of-transit lightcurve data points for a transit detection. The time sampling in the 1999-06 run was too sparse to support transit hunting via the above technique.

In each lightcurve the highest value of  $S_{\text{tra}}^2$  on each night was identified, and those with  $S_{\text{tra}}^2 \geq 100$  (equivalent to  $S/N \geq 10$ ) were retained for closer examination. Table 4.1 lists for each chip the number of raw candidate transits thereby selected over the two runs. Despite the high signal-to-noise threshold for detection, 2182 raw transit candidates were found. A careful visual inspection of the corresponding lightcurves lead us to reject the majority of these based on a number of criteria. The majority of the raw transit candidates (61.8%) were rejected because they appeared to represent a single much fainter data point resulting from a “bad” section in one of the difference images. Such cases were readily identifiable because the lightcurves of many stars triggered a transit detection at the same epoch. A large number of variable stars were picked up ( $\sim 100$  lightcurves  $\equiv 19.5\%$  of the raw transit candidates), which we plan to present in a forthcoming paper. Lightcurves showing eclipses with clearly different depths were also assigned as variable stars since a stellar binary is indicated in this case.

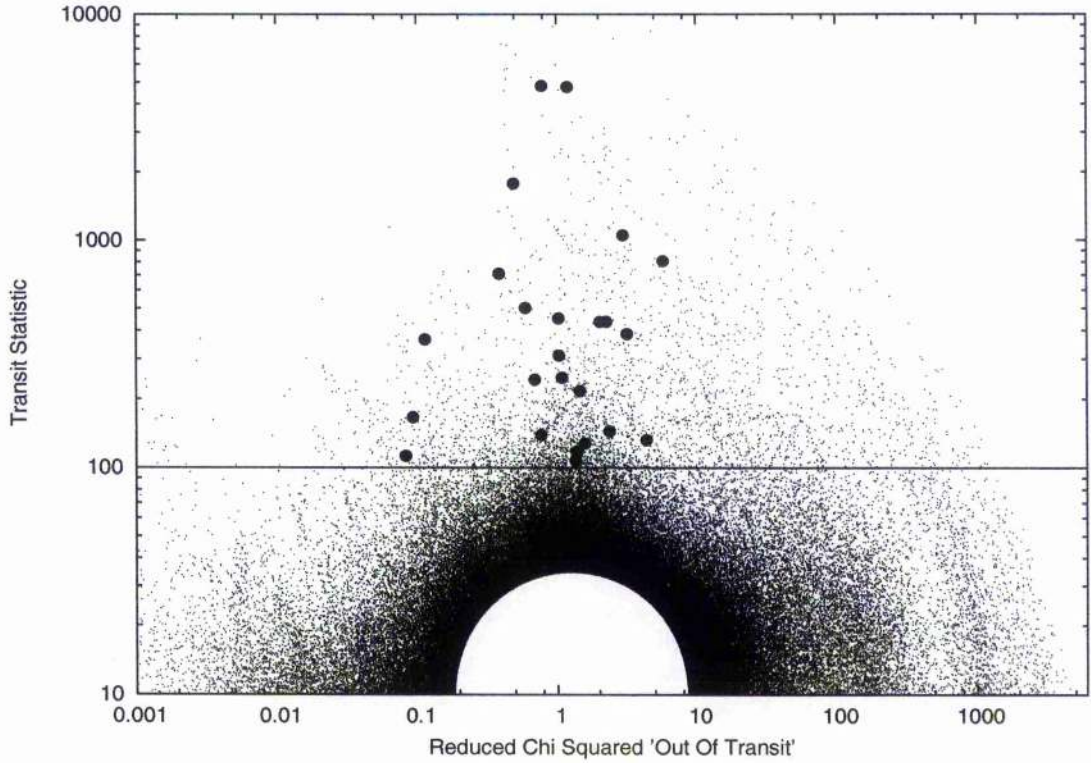


Figure 4.2: The transit statistic  $S_{\text{tra}}^2$  against the out-of-transit reduced chi squared  $\chi_{\text{out}}^2 / (N_{\text{out}} - 1)$ . The initial transit detection threshold shown by the horizontal line is set at  $S_{\text{tra}}^2 = 100$ . The strongest transits detected in the lightcurves of the 24 stars that survived subsequent data quality tests are plotted as solid circles. The blank semicircular region is saturated with test points.

For the remaining raw transit candidates we examined the star on the reference frame. This revealed that many of the remaining transit signatures were caused by the following (in order of most common occurrence):

1. Image defects detected as “stars” (4.5%).
2. Stars lying on or close to image defects, bad columns and/or saturation spikes (4.0%).
3. Stars close to saturation (3.7%).
4. Very closely blended stars (2.1%).
5. Stars close to the edge of the CCD (0.8%).

The reference image for each chip contained a large number of saturated stars along with

large saturation spikes which unfortunately increased the incidence of such false alarms.

For the transit candidate lightcurves that survived to this point, we checked the difference images for the night(s) of the suspected transit(s) by constructing a difference image movie. This revealed that a handful of the candidates (0.6%) were the result of a consecutive set of poor subtractions at the star position. The other candidates clearly showed a flat difference image followed by a growing and then diminishing “dimple”, indicating a drop and then recovery in the brightness of the star.

We discuss in Section 4.2 below the 24 transit candidate lightcurves that passed all of the data quality tests outlined above. Reference to a transit candidate from now on refers only to one of these transit candidate lightcurves. Figure 4.2 shows all tests for which  $S_{\text{tra}}^2 \geq 10$  and highlights the eclipse with the greatest value of  $S_{\text{tra}}^2$  for each transit candidate. Table 4.2 details the number of fully and partially observed eclipses that are present for each transit candidate and how these eclipses are distributed between the three runs. Table 4.2 also lists the J2000.0 celestial coordinates for each transit candidate.

## 4.2 Transit Candidates

### 4.2.1 Theoretical Models

The lightcurves of the 24 transit candidates selected in Section 4.1 were modelled as a star and planet system in the following way. We assume spherical stars, a luminous primary of radius  $R_*$  and a dark massless companion of radius  $R_c$  in a circular orbit with radius  $a$  and period  $P$  inclined by the inclination  $i$  relative to our line of sight. The time  $t_0$  is the time of mid-eclipse of the primary by the companion. Since we already know  $R_*$ , the parameters that need to be constrained for such a system are  $P$ ,  $t_0$ ,  $i$ ,  $R_c$  and a constant magnitude  $m_0$ . Periodic variations in the apparent brightness of the star were also accounted for in



three different ways, leading to three competing planetary transit models:

$$\begin{aligned}
 f_1(t) &= f_0(1 - f_c(t)) && \text{Model 1} \\
 f_2(t) &= f_1(t) \left( 1 + A \sin \left( \frac{2\pi(t - t_0)}{P_{\text{var}}} + \phi \right) \right) && \text{Model 2} \\
 f_3(t) &= f_1(t) \left( 1 - C_e \cos \left( \frac{4\pi(t - t_0)}{P} \right) \right. && \text{Model 3} \\
 &\quad \left. - C_h \cos \left( \frac{2\pi(t - t_0)}{P} \right) \right) && \\
 \end{aligned}
 \tag{4.2}$$

The function  $f_n(t)$  is the predicted stellar flux at time  $t$  for Model  $n$ ,  $f_0$  is a constant flux value and  $f_c(t)$  is the fraction of the total stellar flux obscured by the companion at time  $t$ . We set the linear limb darkening coefficient to  $u = 0.5$  and calculate the function  $f_1(t)/f_0$  using the program *transitcurve.pro* (see Section 2.4.3).

Model 1 is therefore appropriate for a star with a constant brightness. Model 2 incorporates sinusoidal stellar flux variations of semi-amplitude  $A$  and phase  $\phi$  which do not necessarily have the same period as the orbital period of the companion. Such variations may be present for stars with a lot of star spot activity. Model 3 incorporates stellar flux variations due to two effects. The first effect, modelled by the  $C_e$  cosine term, is due to the star being tidally distorted into an ellipsoidal shape by the companion and rotationally synchronised. The value of  $C_e$  quantifies the semi-amplitude of such ellipsoidal flux variations. The second effect, modelled by the  $C_h$  cosine term, is due to heating on one side of the companion caused by irradiation by the star. The value of  $C_h$  quantifies the semi-amplitude of the heating term.

#### 4.2.2 Lightcurve Modelling Procedure

The aim of the lightcurve modelling procedure is to rule out as far as possible the transiting planet model as presented in Section 4.2.1. For nearly half of the transit candidate lightcurves, out-of-eclipse variations were present, and in many cases these variations changed in amplitude and/or phase between the 1999-06, 1999-07 and 2000-09 runs. Hence

it was often necessary to consider each run separately for some or all of the following analysis. Tables 4.3 and 4.4 show the results of the fits of the transiting planet model to the lightcurves. The tables give either a single row per star for fits to all three runs, or three rows per star when the three runs were fitted separately, in which case the first, second and third rows apply to the 1999-06, 1999-07 and 2000-09 runs respectively.

Table 4.2: Star and lightcurve properties for the transit candidates. Star numbers in the ranges 1-19999, 20000-39999, 40000-59999 and 60000-79999 correspond to the chips 1, 2, 3 and 4 respectively. The number in brackets indicates the uncertainty on the last decimal place. Columns 2 and 3 are calibrated  $r'$  and  $r' - z'$  using the shifts in Table 3.5. Columns 6-8 indicate the number of fully and partially observed eclipses ( $f, p$ ) that are present in the lightcurve of the corresponding run. The transit candidates are classified into the following categories: E = Eclipsing binary, EA = Algol type eclipsing binary, RS = RS Canum Venaticorum type eclipsing binary, CV = Cataclysmic variable, INT-7789-TR = Transit candidate that warrants further observations.

Star No.	$r'$ (mag)	$r' - z'$ (mag)	Mass ( $M_{\odot}$ )	Radius ( $R_{\odot}$ )	Distance (pc)	1999-06 $f, p$	1999-07 $f, p$	2000-09 $f, p$	RA (J2000.0)	Dec (J2000.0)	Variable Class
711	20.619	1.133(18) <sup>a</sup>	0.620(9)	0.586(10)	2378(133)	1,1	2,2	5,1	23 58 39.38	+56 36 44.7	RS or CV
791	18.077	0.569(17)	0.949(33)	0.947(42)	2908(13)	0,0	0,1	0,0	23 58 36.75	+56 26 56.1	E
1031	19.140	0.538(9)	1.028(22)	1.062(35)	5762(11)	0,0	0,1	0,2	23 58 33.88	+56 37 04.9	E
1262	18.900	0.788(9) <sup>a</sup>	0.740(3)	0.707(3)	2037(39)	0,5	2,3	5,2	23 58 29.19	+56 32 42.7	RS
5917	16.724	0.55(8) <sup>b</sup>	0.94(17)	0.93(25)	1589(17)	0,1	1,2	0,0	23 57 12.86	+56 31 26.5	EA
5974	19.346	0.606(12) <sup>a</sup>	0.897(23)	0.882(29)	4346(31)	0,0	0,1	0,1	23 57 11.92	+56 31 24.9	RS
6995	21.120	1.853(19)	0.239(6)	0.245(4)	752(4)	0,1	1,0	1,1	23 56 57.23	+56 34 03.4	E
7628	19.420	1.298(9) <sup>a</sup>	0.544(5)	0.507(4)	1043(14)	1,0	2,0	2,0	23 56 47.64	+56 36 28.7	RS
7695	19.447	0.999(10)	0.666(4)	0.635(4)	1799(4)	0,0	0,0	0,1	23 56 46.79	+56 36 13.8	E
21790	19.728	0.562(13) <sup>a</sup>	0.980(23)	0.988(34)	6619(59)	1,1	3,0	2,1	23 55 59.17	+56 45 14.1	EA
22688	19.981	0.925(12) <sup>a</sup>	0.701(4)	0.671(4)	2647(51)	1,2	3,1	5,2	23 55 18.41	+56 43 14.3	RS
22738	17.650	0.64(9) <sup>a</sup>	0.81(12)	0.78(14)	1578(10)	0,0	0,1	0,2	23 56 01.68	+56 43 08.3	RS
23979	20.088	0.773(18) <sup>a</sup>	0.751(5)	0.716(5)	3564(20)	0,0	1,1	1,1	23 55 58.93	+56 40 29.6	E
24512	17.338	0.695(45)	0.759(23)	0.723(24)	1168(2)	0,1	0,1	0,1	23 55 42.99	+56 39 14.9	E
45134	20.703	0.632(32)	0.87(5)	0.84(6)	7291(30)	0,0	0,1	1,0	23 57 45.06	+56 55 36.6	INT-7789-TR-1
46271	20.323	1.025(19) <sup>b</sup>	0.662(8)	0.630(8)	2545(79)	0,0	1,1	3,2	23 57 29.99	+56 57 34.3	RS
46691	18.024	0.467(8)	1.195(22)	1.348(37)	5348(7)	0,0	0,0	1,0	23 57 24.52	+56 55 17.6	INT-7789-TR-2
47171	17.571	0.427(10) <sup>a</sup>	1.33(4)	1.57(7)	5599(85)	0,3	1,2	3,1	23 57 18.04	+56 51 12.0	EA
49512	19.553	0.971(9)	0.679(4)	0.649(4)	1995(5)	1,0	0,0	2,1	23 56 47.11	+56 51 10.2	INT-7789-TR-3
50313	17.368	0.62(7)	0.82(10)	0.79(12)	1477(2)	0,1	0,1	0,2	23 56 36.67	+56 52 43.4	RS
61876	20.865	0.560(37)	0.98(8)	1.00(13)	11312(52)	0,0	1,0	1,0	23 58 27.34	+56 46 36.4	E
62983	20.402	0.616(24)	0.89(5)	0.87(6)	6784(22)	0,0	0,0	3,2	23 58 13.37	+56 45 36.1	E
64804	20.080	1.516(9)	0.399(6)	0.379(5)	526(2)	0,1	0,0	1,0	23 57 51.16	+56 42 03.2	E
73852	18.984	0.797(18) <sup>a</sup>	0.743(6)	0.710(5)	3423(84)	0,0	1,0	2,1	23 56 11.77	+56 45 55.6	RS

<sup>a</sup> Colour corrected using the lightcurve model. <sup>b</sup> Colour corrected using interpolation of the lightcurve.

The procedure used to model the transit candidate lightcurves is as follows. Each lightcurve was inspected and the best-defined eclipse chosen. The lightcurve data for the night on which this eclipse occurred were fitted with Model 1 keeping the inclination fixed at  $90.0^\circ$  (we call this a central transit fit). This fit determines a time of mid-transit  $t_0$ , a constant magnitude  $m_0$ , a companion radius  $R_{c,\min}$  and a transit duration  $\Delta t$ . The value of  $R_{c,\min}$  is the minimum radius of a companion given the eclipse profile, since at lower inclination values the same size companion obscures a smaller fraction of the total stellar flux due to limb darkening effects and the possibility that the eclipse is grazing instead of annular. Hence, at lower inclinations a larger companion radius is required to account for the observed eclipse depth. If no other eclipses were present in the lightcurve then the value of  $R_{c,\min}$  is presented in Tables 4.3 and 4.4 as the value of  $R_c$  for a fixed inclination of  $90.0^\circ$ , along with the fitted values of  $t_0$ ,  $m_0$  and  $\Delta t$ . Out-of-eclipse variations, if present, were also fitted by including the relevant parameters in the central transit fit.

When more than one eclipse was evident in the lightcurve, we constructed an eclipse periodogram in order to determine the orbital period  $P$ . Folding the lightcurve around the time of mid-transit  $t_0$  using the period  $P$ , we then calculate the chi squared  $\chi^2_{\text{const}}$  of a constant fit to the folded lightcurve and the chi squared  $\chi^2_{\text{central}}$  of the previously determined central transit model optimally scaled to fit the folded lightcurve at time  $t_0$ . We thereby define:

$$\Delta\chi^2 = \chi^2_{\text{const}} - \chi^2_{\text{central}} \quad (4.3)$$

representing the improvement in  $\chi^2$  at period  $P$  for the scaled central transit model as opposed to the constant model. The period producing the highest value of  $\Delta\chi^2$  is adopted as the best fit. Since the ratio of the eclipse duration to the orbital period is generally small, and with many orbital cycles elapsing over the span of our observations, the above method can lead to very accurate period determinations. We obtained a conservative estimate for the uncertainty in  $P$  by fitting a gaussian to the local peak in  $\Delta\chi^2$  and calculating the half width half maximum.

On folding the lightcurve on the period  $P$  determined by the eclipse periodogram, it sometimes became obvious that the true period was  $2P$ , in which case the value of  $P$  was



updated. If out-of-eclipse variations were present, then the folded lightcurve would reveal whether or not  $P_{\text{var}} = P$ . If it was the case that  $P_{\text{var}} \neq P$ , then a sine periodogram was constructed in order to determine  $P_{\text{var}}$ , using the lightcurve with the eclipses masked out. The method used is the same as that for an eclipse periodogram, except that a sine curve and a cosine curve both of period  $P$  are optimally scaled to the folded lightcurve instead of the central transit model, and a value of chi squared  $\chi^2_{\text{sine}}$  calculated. The value of  $P_{\text{var}}$  so determined was usually not very accurate due to the sine curve cycle being the same length as the period being tested, and due to the fact that a single run was often used instead of all three as a result of the changing amplitude and/or phase of the out-of-eclipse variations.

Table 4.3: Lightcurve and companion properties for the transit candidates as obtained from the fit of the transiting planet model to the lightcurve data. See Section 4.2.1 for the definitions of the parameters in the column headers. Column 4 is calibrated  $r'$  magnitude using the shifts in Table 3.5. The column marked  $\chi^2/N$  is the reduced chi squared of the fit.

Star No.	$t_0 - 2451000.0$ (d)	$\left[ \begin{array}{c} P_{\text{var}} \\ P \end{array} \right]$ (d)	$m_0$ ( $r'$ mag)	$\Delta m$ (mag)	$\Delta t$ (h)	$i$ ( $^\circ$ )	$R_c$ ( $R_\odot$ )	A or $C_e$	$\phi$ (rad) or $C_h$	$\chi^2/N$
711	808.4805(11)	$\left[ \begin{array}{c} 0.450285(20) \\ 0.450285(20) \end{array} \right]$	$\left\{ \begin{array}{c} 20.668 \\ 20.666 \\ 20.619 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.40 \\ 0.31 \\ 0.24 \end{array} \right\}$	$\left\{ \begin{array}{c} 1.47 \\ 1.47 \\ 1.33 \end{array} \right\}$	$\left\{ \begin{array}{c} 81.1(3.4) \\ 76.3(3.6) \\ 82.2(1.0) \end{array} \right\}$	$\left\{ \begin{array}{c} 0.33(8) \\ 0.40(6) \\ 0.25(1) \end{array} \right\}$	$\left\{ \begin{array}{c} A = 0.0323 \\ A = 0.0395 \\ A = 0.1008 \end{array} \right\}$	$\left\{ \begin{array}{c} \phi = 3.292 \\ \phi = 3.015 \\ \phi = 3.484 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.97 \\ 1.06 \\ 2.54 \end{array} \right\}$
791	384.481(4)	$\left[ \begin{array}{c} - \\ - \end{array} \right]$	18.077	0.25	2.50	90.0*	0.39(2)	—	—	3.57
1031	801.584(9)	$\left[ \begin{array}{c} - \\ 3.6216(53)^\dagger \end{array} \right]$	19.140	0.12	5.20	90.0*	>0.32	—	—	2.31
1262	805.5611(7)	$\left[ \begin{array}{c} 0.856336(40) \\ 0.856336(40) \end{array} \right]$	$\left\{ \begin{array}{c} 18.937 \\ 18.931 \\ 18.900 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.141 \\ 0.156 \\ 0.093 \end{array} \right\}$	$\left\{ \begin{array}{c} 1.68 \\ 1.81 \\ 1.59 \end{array} \right\}$	$\left\{ \begin{array}{c} 83.1(5) \\ 74.3(2.0) \\ 83.7(9) \end{array} \right\}$	$\left\{ \begin{array}{c} 0.353(10) \\ 0.83(10) \\ 0.276(9) \end{array} \right\}$	$\left\{ \begin{array}{c} A = 0.0116 \\ A = 0.0089 \\ A = 0.0411 \end{array} \right\}$	$\left\{ \begin{array}{c} \phi = 1.524 \\ \phi = 3.011 \\ \phi = 1.512 \end{array} \right\}$	$\left\{ \begin{array}{c} 12.3 \\ 6.86 \\ 5.00 \end{array} \right\}$
5917	391.5539(9)	$\left[ \begin{array}{c} 1.03377(75) \\ 1.03377(75) \end{array} \right]$	$\left\{ \begin{array}{c} 16.741 \\ 16.724 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.14 \\ 0.16 \end{array} \right\}$	$\left\{ \begin{array}{c} 2.34 \\ 2.21 \end{array} \right\}$	$\left\{ \begin{array}{c} 89.8(9.8) \\ 83.0(2) \end{array} \right\}$	$\left\{ \begin{array}{c} 0.29(22) \\ 0.33(9) \end{array} \right\}$	$\left\{ \begin{array}{c} C_e = 0.0021 \\ C_e = 0.0010 \end{array} \right\}$	$\left\{ \begin{array}{c} C_h = 0.0085 \\ C_h = 0.0138 \end{array} \right\}$	$\left\{ \begin{array}{c} 2.74 \\ 4.56 \end{array} \right\}$
5974	798.5967(14)	$\left[ \begin{array}{c} 7.5(1.9) \\ - \end{array} \right]$	19.346	0.055	3.24	90.0*	>0.18	$A = 0.0108$	$\phi = 2.745$	1.32
6995	798.4653(5)	$\left[ \begin{array}{c} - \\ 1.310743(98) \end{array} \right]$	21.120	0.39	1.28	84.8(2)	0.245(12)	—	—	1.95
7628	808.6076(7)	$\left[ \begin{array}{c} 1.759555(93) \\ 1.759555(93) \end{array} \right]$	$\left\{ \begin{array}{c} 19.378 \\ 19.373 \\ 19.420 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.15 \\ 0.15 \\ 0.16 \end{array} \right\}$	$\left\{ \begin{array}{c} 1.60 \\ 1.62 \\ 1.70 \end{array} \right\}$	$\left\{ \begin{array}{c} 85.0(3.0) \\ 85.2(1.7) \\ 83.7(6) \end{array} \right\}$	$\left\{ \begin{array}{c} 0.23(19) \\ 0.23(11) \\ 0.33(4) \end{array} \right\}$	$\left\{ \begin{array}{c} A = 0.0354 \\ A = 0.0364 \\ A = 0.0224 \end{array} \right\}$	$\left\{ \begin{array}{c} \phi = 4.422 \\ \phi = 4.289 \\ \phi = 1.275 \end{array} \right\}$	$\left\{ \begin{array}{c} 1.94 \\ 1.80 \\ 2.27 \end{array} \right\}$
7695	805.4288(19)	$\left[ \begin{array}{c} - \\ - \end{array} \right]$	19.447	0.35	4.61	90.0*	0.309(2)	—	—	2.84
21790	801.5709(8)	$\left[ \begin{array}{c} 0.800610(96) \\ 0.800610(96) \end{array} \right]$	19.728	0.081	1.81	78.7(6)	0.261(19)	$C_e = 0.0103$	$C_h = 0.0016$	1.23
22688	800.6104(8)	$\left[ \begin{array}{c} 0.904404(53) \\ 0.904404(53) \end{array} \right]$	$\left\{ \begin{array}{c} 20.051 \\ 20.003 \\ 19.981 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.15 \\ 0.18 \\ 0.18 \end{array} \right\}$	$\left\{ \begin{array}{c} 1.83 \\ 1.93 \\ 1.92 \end{array} \right\}$	$\left\{ \begin{array}{c} 73.2(3.1) \\ 73.9(1.6) \\ 73.8(9) \end{array} \right\}$	$\left\{ \begin{array}{c} 0.67(17) \\ 0.67(8) \\ 0.67(5) \end{array} \right\}$	$\left\{ \begin{array}{c} A = 0.0073 \\ A = 0.0255 \\ A = 0.0357 \end{array} \right\}$	$\left\{ \begin{array}{c} \phi = 0.299 \\ \phi = 6.168 \\ \phi = 0.539 \end{array} \right\}$	$\left\{ \begin{array}{c} 2.60 \\ 4.36 \\ 2.65 \end{array} \right\}$
22738	800.6278(12)	$\left[ \begin{array}{c} 2.769(50) \\ 2.769(50) \end{array} \right]$	$\left\{ \begin{array}{c} - \\ - \end{array} \right\}$	$\left\{ \begin{array}{c} - \\ - \end{array} \right\}$	$\left\{ \begin{array}{c} - \\ - \end{array} \right\}$	$\left\{ \begin{array}{c} - \\ - \end{array} \right\}$	$\left\{ \begin{array}{c} - \\ - \end{array} \right\}$	$\left\{ \begin{array}{c} - \\ - \end{array} \right\}$	$\left\{ \begin{array}{c} - \\ - \end{array} \right\}$	$\left\{ \begin{array}{c} - \\ - \end{array} \right\}$
23979	798.5229(11)	$\left[ \begin{array}{c} - \\ - \end{array} \right]$	20.088	0.56	2.56	89.9(2)	0.27(5)	$A = 0.0078$	$\phi = 2.640$	1.91
24512	387.457(11)	$\left[ \begin{array}{c} - \\ - \end{array} \right]$	17.338	0.11	3.36	90.0*	0.426(3)	—	—	2.05
							>0.20	—	—	3.38

<sup>†</sup>Uncertain period (see Section 4.2.3). \*Inclination fixed and  $\Delta t$  fitted (central transit fit).

Table 4.4: Lightcurve and companion properties for the transit candidates (continued) as obtained from the fit of the transiting planet model to the lightcurve data. See Section 4.2.1 for the definitions of the parameters in the column headers. Column 4 is calibrated  $r'$  magnitude using the shifts in Table 3.5. The column marked  $\chi^2/N$  is the reduced chi squared of the fit.

Star No.	$t_0 - 2451000.0$ (d)	$\left[ \begin{array}{c} P_{\text{var}} \\ P \end{array} \right]$ (d)	$m_0$ ( $r'$ mag)	$\Delta m$ (mag)	$\Delta t$ (h)	$i$ ( $^\circ$ )	$R_c$ ( $R_\odot$ )	A or $C_e$	$\phi$ (rad) or $C_h$	$\chi^2/N$
45134 $^\circ$	806.4638(21)	$\left[ \begin{array}{c} 58(12)^\dagger \\ 0.902117(61) \\ 0.902117(61) \end{array} \right]$	20.703	0.072	5.96	89.1(3)	0.42(22)	—	—	1.65
46271	801.1921(8)	$\left[ \begin{array}{c} 0.902117(61) \\ 0.902117(61) \end{array} \right]$	$\left\{ \begin{array}{c} 20.324 \\ 20.323 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.18 \\ 0.22 \end{array} \right\}$	$\left\{ \begin{array}{c} 1.82 \\ 1.91 \end{array} \right\}$	$\left\{ \begin{array}{c} 75(11) \\ 75.2(2) \end{array} \right\}$	$\left\{ \begin{array}{c} 0.63(41) \\ 0.628(18) \end{array} \right\}$	$\left\{ \begin{array}{c} A = 0.0238 \\ A = 0.0647 \end{array} \right\}$	$\left\{ \begin{array}{c} \phi = 3.354 \\ \phi = 2.321 \end{array} \right\}$	$\left\{ \begin{array}{c} 2.30 \\ 2.01 \end{array} \right\}$
46691 $^\circ$	805.5405(9)	$\left[ \begin{array}{c} 1.8(1.3)^\dagger \\ 0.856026(73) \\ 0.856026(73) \end{array} \right]$	18.024	0.021	2.50	81.9(1.6)	0.185(9)	—	—	1.44
47171	805.5715(12)	$\left[ \begin{array}{c} 0.856026(73) \\ 0.856026(73) \end{array} \right]$	$\left\{ \begin{array}{c} 17.590 \\ 17.613 \\ 17.571 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.069 \\ 0.085 \\ 0.130 \end{array} \right\}$	$\left\{ \begin{array}{c} 2.52 \\ 2.76 \\ 2.97 \end{array} \right\}$	$\left\{ \begin{array}{c} 69.4(3.3) \\ 67.6(5.7) \\ 72.6(9) \end{array} \right\}$	$\left\{ \begin{array}{c} 0.51(18) \\ 0.67(28) \\ 0.58(6) \end{array} \right\}$	$\left\{ \begin{array}{c} C_e = 0.0109 \\ C_e = 0.0088 \\ C_e = 0.0172 \end{array} \right\}$	$\left\{ \begin{array}{c} C_h = 0.0097 \\ C_h = 0.0000 \\ C_h = 0.0046 \end{array} \right\}$	$\left\{ \begin{array}{c} 3.07 \\ 2.85 \\ 1.09 \end{array} \right\}$
49512 $^\circ$	800.6303(7)	$\left[ \begin{array}{c} 1.2431(58) \end{array} \right]$	$\left\{ \begin{array}{c} 19.553 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.070 \end{array} \right\}$	$\left\{ \begin{array}{c} 1.72 \end{array} \right\}$	$\left\{ \begin{array}{c} 87.1(1.2) \end{array} \right\}$	$\left\{ \begin{array}{c} 0.151(7) \end{array} \right\}$	—	—	1.28
50313	798.6707(15)	$\left[ \begin{array}{c} 5.664(16) \end{array} \right]$	$\left\{ \begin{array}{c} 17.368 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.16 \end{array} \right\}$	$\left\{ \begin{array}{c} 3.68 \end{array} \right\}$	$\left\{ \begin{array}{c} 84.6(1) \end{array} \right\}$	$\left\{ \begin{array}{c} 0.79(12) \end{array} \right\}$	—	—	2.83
61876	807.5492(8)	$\left[ \begin{array}{c} - \end{array} \right]$	20.865	0.13	2.03	90.0*	0.30(4)	—	—	0.74
62983	806.5823(9)	$\left[ \begin{array}{c} 1.074(20) \end{array} \right]$	$\left\{ \begin{array}{c} 20.402 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.18 \end{array} \right\}$	$\left\{ \begin{array}{c} 2.12 \end{array} \right\}$	$\left\{ \begin{array}{c} 81.3(1.1) \end{array} \right\}$	$\left\{ \begin{array}{c} 0.38(7) \end{array} \right\}$	—	—	1.80
64804	801.6078(4)	$\left[ \begin{array}{c} - \end{array} \right]$	18.984	0.19	0.98	90.0*	0.141(2)	—	—	1.35
73852	806.4952(10)	$\left[ \begin{array}{c} 1.531(18) \\ 1.531(18) \end{array} \right]$	$\left\{ \begin{array}{c} 20.080 \end{array} \right\}$	$\left\{ \begin{array}{c} 0.40 \end{array} \right\}$	$\left\{ \begin{array}{c} 2.43 \end{array} \right\}$	$\left\{ \begin{array}{c} 85.9(3) \end{array} \right\}$	$\left\{ \begin{array}{c} 0.390(11) \end{array} \right\}$	$\left\{ \begin{array}{c} A = 0.0447 \end{array} \right\}$	$\left\{ \begin{array}{c} \phi = 4.036 \end{array} \right\}$	$\left\{ \begin{array}{c} 2.09 \end{array} \right\}$

\*Inclination fixed and  $\Delta t$  fitted (central transit fit).  $^\dagger$ Period and inclination predicted, and  $\Delta t$  fitted (see Sections 4.2.8 & 4.2.9).  $^\circ$ Transit candidates that warrant further observations (45134 $\equiv$ INT-7789-TR-1, 46691 $\equiv$ INT-7789-TR-2 and 49512 $\equiv$ INT-7789-TR-3).

The final stage in the lightcurve modelling procedure was to fit the appropriate model to the lightcurve (we call this a full transit fit). For lightcurves where out-of-transit variations were present, the choice between Model 2 and Model 3 was made by adopting the model that produced the smallest  $\chi^2$  when the component of the model accounting for the stellar brightness variations was fitted to the lightcurve with the eclipses masked out. The transiting planet model was fitted by calculating the  $\chi^2$  of the chosen model for a grid in orbital inclination  $i$  and companion radius  $R_c$ , optimising the remaining parameters of the model at each grid point, and using the values of  $t_0$ ,  $P$  and  $P_{\text{var}}$  determined earlier as initial values. The value of  $P$  was not adjusted in the fit if it had been determined from an eclipse periodogram using lightcurve data from all three runs and if the lightcurve data being fitted were from a single run (since this value of  $P$  is already more accurate than that which can be determined from a single run). The minimum of the constructed  $\chi^2$  surface indicates the best fit values for  $i$  and  $R_c$ . From the fitted parameters an updated transit duration  $\Delta t$  was calculated. Confidence regions at the 1, 2 and  $3\sigma$  levels were calculated by constructing contours at  $\chi^2 = \chi^2_{\text{min}} + 2.30$ ,  $\chi^2_{\text{min}} + 6.17$  and  $\chi^2_{\text{min}} + 11.8$  where  $\chi^2_{\text{min}}$  is the minimum  $\chi^2$  value. The uncertainties on  $i$  and  $R_c$  were obtained by projecting the  $\chi^2_{\text{min}} + 1.0$  confidence region onto each parameter axis. The uncertainties so derived on these parameters can be quite large due to the strong correlation between  $i$  and  $R_c$ . It is possible to choose a family of values for  $i$  and  $R_c$  that produce approximately the same eclipse depth and duration, where only the shape of the eclipse differs slightly. Our observations were not always of sufficient accuracy and/or frequency to strongly constrain the eclipse shape and hence the confidence regions can be quite extensive.

At this stage, various arguments can be invoked to rule out the transiting planet model. Firstly, if on folding the lightcurve on the orbital period  $P$  it becomes apparent that there are eclipses of different depth, then we classify the system as an eclipsing stellar binary. Similarly, if out-of-eclipse variations are present and the lightcurve data are best modelled by Model 3 (ellipsoidal variations and/or heating effects), then we also classify the system as an eclipsing stellar binary. Finally, if the companion radius (or minimum companion radius) is greater than  $0.2R_{\odot}$ , then the companion is most likely to be a star based on the known radii of hot (and cold) Jupiters to date.

In the following sections, the transit candidates have been organised into groups depending on the lightcurve properties, and analysed according to the above methodology. For brevity, the following labelling format has been adopted for the plots in these sections:

<Star No.> - <Plot Code> - <Run(s) To Which The Plot Applies>

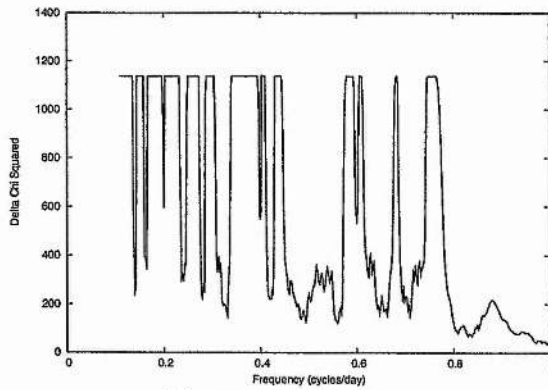
The plot codes are as follows:

1. EP - Eclipse periodogram
2. SP - Sine periodogram
3. CM - Chi squared contour map showing the best fit solution with a cross and the 1, 2 and  $3\sigma$  confidence regions with solid, dashed and shorter dashed lines respectively. Annular, grazing and no eclipse regions are separated by thick solid lines.
4. L - Lightcurve
5. PL - Phased lightcurve
6. CPL - Close up of phased lightcurve
7. BL - Binned lightcurve
8. CTFIT - Central transit fit
9. FTFIT - Full transit fit

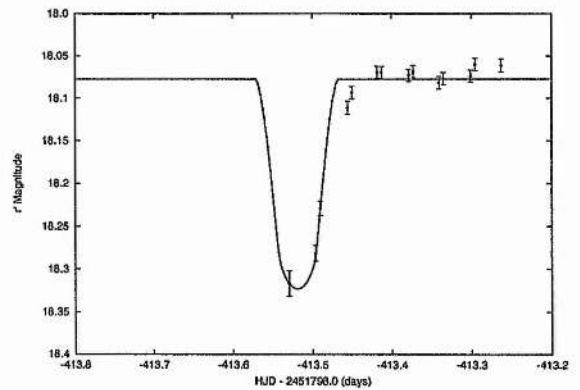
#### 4.2.3 Eclipsing Binaries With Undetermined Periods

In this section we present 7 transit candidates for which we were unable to determine a period, although we were able to classify them as eclipsing binaries. The reason for not being able to determine the period was due to either the presence of only one fully/partially observed eclipse in the lightcurve and/or cycle ambiguity between eclipses.

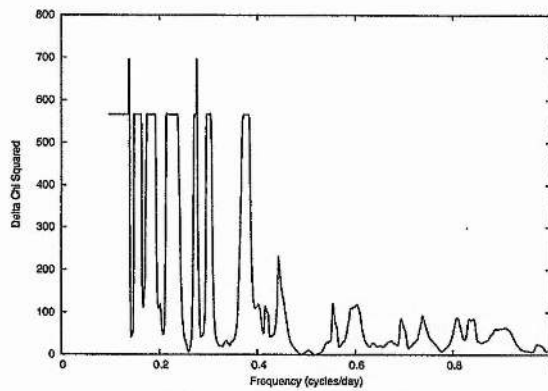
**Star 791** shows a partially observed, poorly sampled, eclipse of depth 0.25 mag in the 1999-07 data (Figure 4.3(b)) that hardly constrains the eclipse duration. Although the period is unknown, the eclipse periodogram (Figure 4.3(a)) shows that we can rule out all periods shorter than 1.25 d. With  $r' \approx 18.08$  mag and  $r' - i' \approx 0.57$  mag we find that the



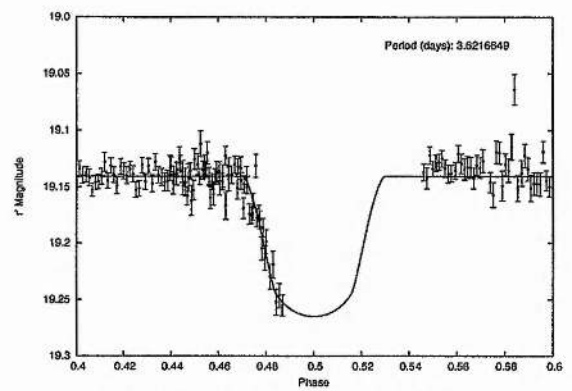
(a) 791 - EP - 1999-07



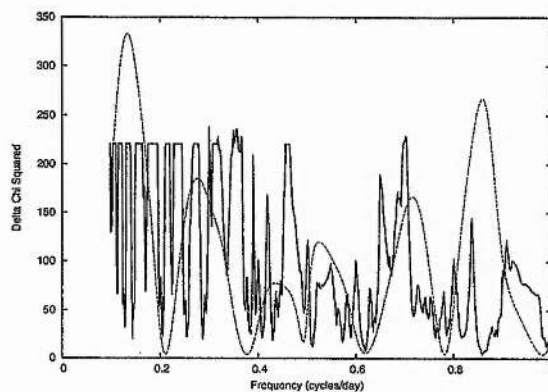
(b) 791 - L &amp; CTFIT - 1999-07 Night 3



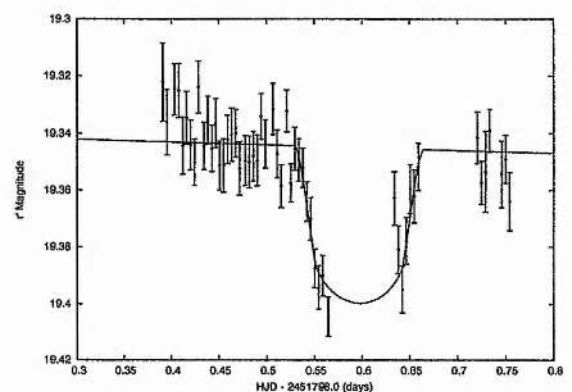
(c) 1031 - EP - 2000-09



(d) 1031 - CPL &amp; CTFIT - 2000-09

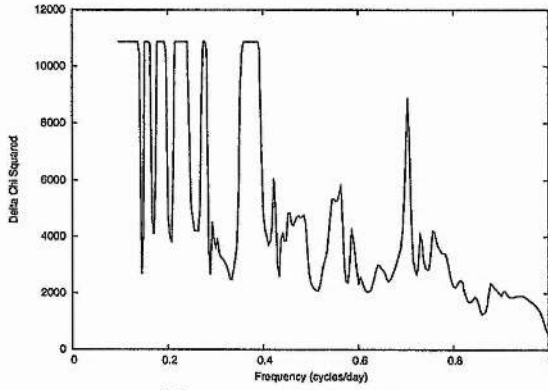


(e) 5974 - EP (continuous line) &amp; SP (dashed line) - 2000-09

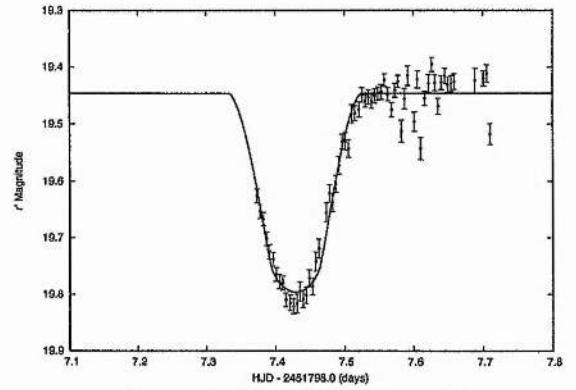


(f) 5974 - L &amp; CTFIT - 2000-09 Night 1

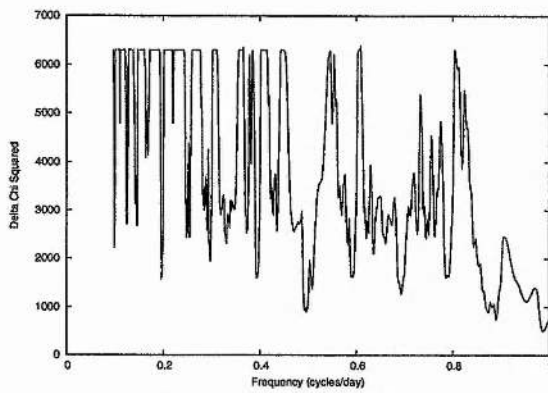
Figure 4.3: Eclipsing binaries with undetermined periods.



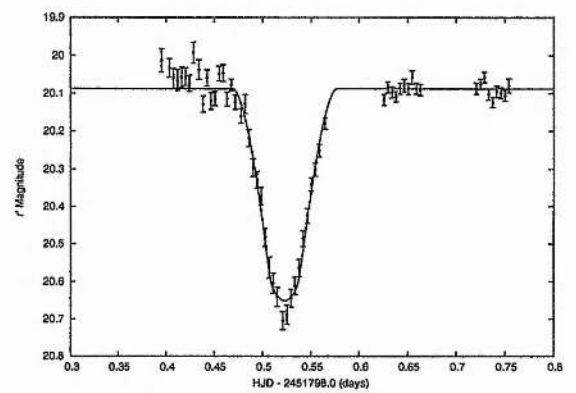
(a) 7695 - EP - 2000-09



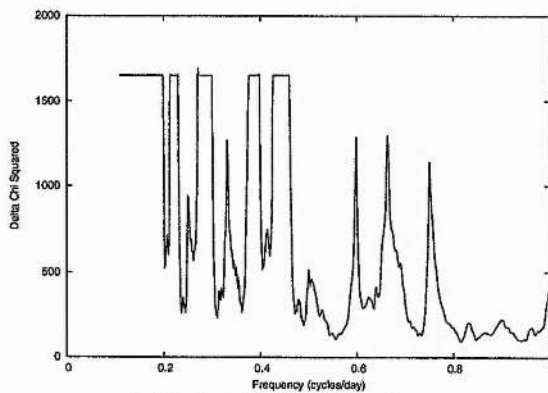
(b) 7695 - L &amp; CTFIT - 2000-09 Night 8



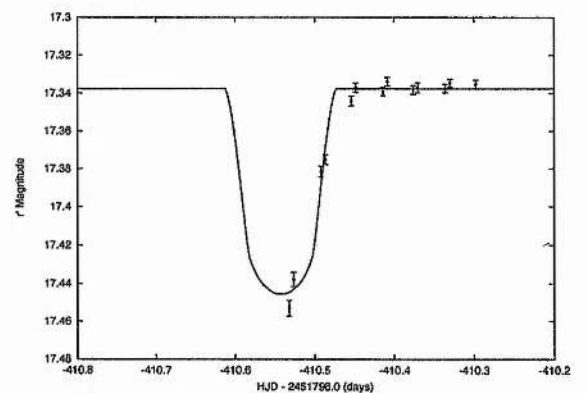
(c) 23979 - EP - 2000-09



(d) 23979 - L &amp; CTFIT - 2000-09 Night 1



(e) 24512 - EP - 1999-07



(f) 24512 - L &amp; CTFIT - 1999-07 Night 6

Figure 4.4: Eclipsing binaries with undetermined periods.



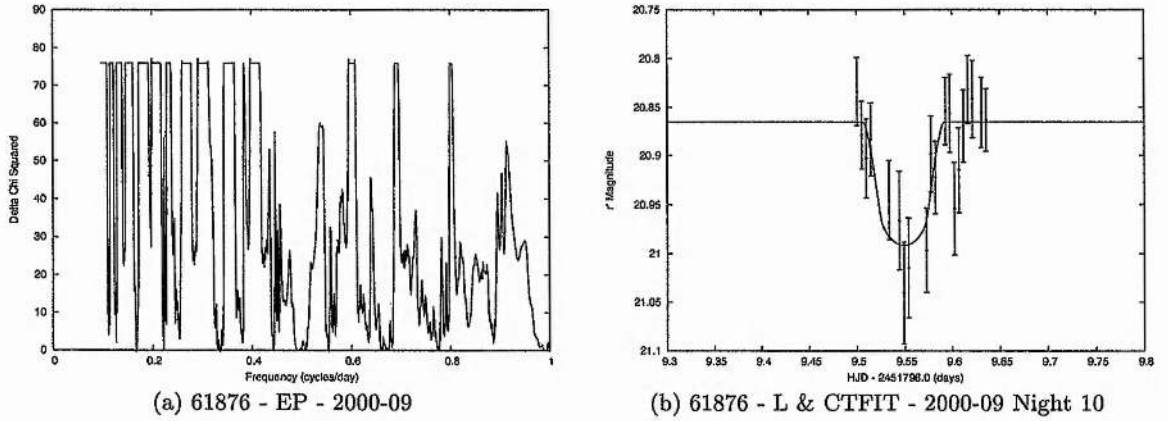


Figure 4.5: Eclipsing binaries with undetermined periods.

primary is a late G star of mass  $0.95M_{\odot}$  that lies at  $d = 2.9\text{kpc}$ , slightly beyond the cluster. We derive a minimum companion radius of  $0.39R_{\odot}$ .

**Star 1031** shows three partially observed eclipses of which the best occur during the 2000-09 run (Table 4.2). Neither the eclipse depth nor duration are constrained by the observations, although the eclipse periodogram (Figure 4.3(c)) reveals that there are only two possible periods ( $3.6216 \pm 0.0053$  d and  $7.233 \pm 0.010$  d). We plot the 2000-09 lightcurve folded on the shorter period in Figure 4.3(d). The star has  $r' \approx 19.14$  mag and  $r' - i' \approx 0.54$  mag giving a primary star mass of  $1.03M_{\odot}$  and suggesting that it is a G star similar to our Sun that lies beyond the cluster at  $d = 5.8\text{kpc}$ . We derive a robust minimum companion radius of  $0.32R_{\odot}$ .

**Star 5974** exhibits two partially observed eclipses of duration 3.2 h with a poorly defined depth. Lightcurve modulations of amplitude 0.02 mag and period 7.5 d that are out of phase with the eclipses suggest the presence of star spots. The 2000-09 eclipse is shown in Figure 4.3(f). The eclipse periodogram (Figure 4.3(e)) rules out periods shorter than 1.3 d. With  $r' \approx 19.35$  mag and  $r' - i' \approx 0.61$  mag we derive a primary mass of  $0.90M_{\odot}$ , spectral type G6V and distance of 4.3kpc, placing the system beyond the cluster. We derive a robust minimum companion radius of  $0.18R_{\odot}$ , most likely an under estimate due to the possibility that the eclipse is deeper than the fit shown in Figure 4.3(f).

**Star 7695** shows a nearly complete, well sampled, 4.6 h eclipse of depth 0.35 mag (Figure 4.4(b)). The relatively long eclipse duration allows all periods shorter than 2.50 d to be ruled out (Figure 4.4(a)). The star has  $r' \approx 19.45$  mag and  $r' - i' \approx 1.0$  mag suggesting a  $0.67M_{\odot}$  K5V primary at  $d = 1.8$  kpc, in front of the cluster. We derive a minimum companion radius of  $0.309R_{\odot}$ .

**Star 23979** exhibits one fully and one partially observed eclipse in each of the 1999-07 and 2000-09 runs. The best defined eclipse (Figure 4.4(d)) reveals an eclipse depth of 0.56 mag and duration 2.6 h. The eclipse periodogram (Figure 4.4(c)) rules out periods shorter than 1.1 d, but otherwise does not suggest a likely period. With  $r' \approx 20.09$  mag and  $r' - i' \approx 0.77$  mag we find that the primary is a  $0.75M_{\odot}$  K2V star at  $d = 3.6$  kpc, beyond the cluster. We derive a minimum companion radius of  $0.426R_{\odot}$ .

**Star 24512** shows a partially observed eclipse in each of the three runs, leaving the eclipse depth and duration unknown. The best eclipse is shown in Figure 4.4(f) which suggests a minimum depth of 0.1 mag. The eclipse periodogram (Figure 4.4(e)) shows that we can rule out periods shorter than 2.0 d. The star has  $r' \approx 17.34$  mag and  $r' - i' \approx 0.70$  mag suggesting another  $0.76M_{\odot}$  K2V primary star that lies at  $d = 1.2$  kpc, half the cluster distance. We derive a robust minimum companion radius of  $0.20R_{\odot}$ .

**Star 61876** exhibits two eclipses of depth 0.13 mag and duration 2.0 h, with the best sampled eclipse shown in Figure 4.5(b). We can rule out periods shorter than 1.15 d from the eclipse periodogram in Figure 4.5(a). With  $r' \approx 20.87$  mag and  $r' - i' \approx 0.56$  mag we derive a primary mass of  $0.98M_{\odot}$ , spectral type G2V like our Sun and distance of 11.3 kpc, well beyond the cluster and beyond the “edge” of the galaxy (see Section 3.4.4). We derive a minimum companion radius of  $0.30R_{\odot}$ .

For each of the above transit candidates except star 5974, the central transit fit yields a minimum companion radius that is greater than  $0.2R_{\odot}$ . This favours a stellar rather than a planetary companion. The lack of out-of-eclipse lightcurve variations leads us to conclude that these are eclipsing binaries. As we have already mentioned for star 5974, the minimum companion radius of  $0.18R_{\odot}$  is most likely an under estimate, and the lightcurve shows sinusoidal out-of-eclipse variations. Therefore we class this system as a RS CVn type

eclipsing binary.

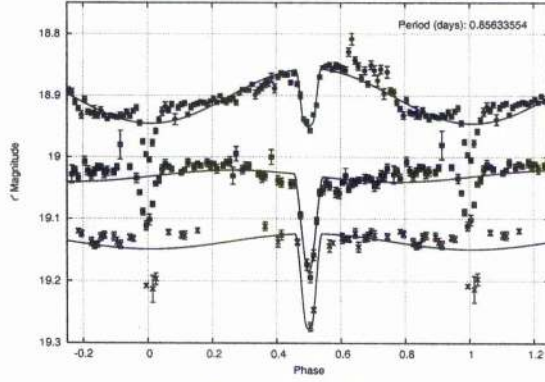
#### 4.2.4 Eclipsing Binaries Exhibiting Secondary Eclipses

In this section we present 4 transit candidates that exhibit secondary eclipses in their lightcurves implying that the companion is luminous. Figure 4.6 shows the folded lightcurve for each transit candidate along with the best fit transiting planet model. The lightcurves from the different runs are offset vertically (in magnitude) from each other in order to highlight any changes in the out-of-eclipse variations.

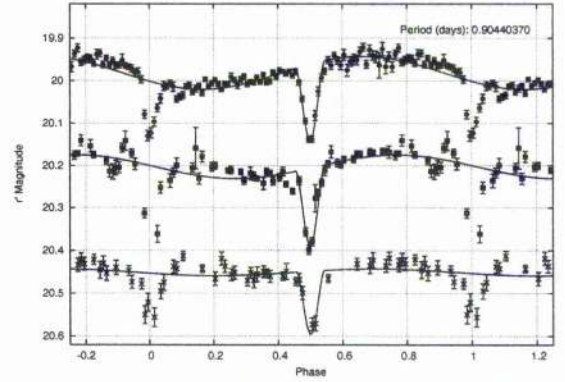
**Star 1262** shows eclipses of different depths ( $\sim 0.05$  mag difference) in the 1999-07 lightcurve data (Figure 4.6(a)), revealing the secondary eclipse, along with lightcurve modulations that change in phase and increase in amplitude from 0.02 mag to 0.08 mag over the three runs. The system has a period of 0.856 d, eclipses of duration 1.6 h and a primary eclipse that decreases in depth from 0.15 to 0.09 mag over the three runs. The changing lightcurve modulations and primary eclipse depth are suggestive of spot activity on the primary star. The star has  $r' \approx 18.90$  and  $r' - i' \approx 0.79$  suggesting a  $0.74M_{\odot}$  K2V primary star that lies at  $d = 2.0$  kpc, slightly in front of the cluster. Our fits to the lightcurves are inconclusive about the size of the secondary star (Table 4.3). We classify this system as a RS CVn type eclipsing binary.

**Star 22688** is a 0.904 d eclipsing binary with eclipses of different depths ( $\sim 0.03$  mag difference - Figure 4.6(b)). The primary eclipse is of depth 0.18 mag and duration 1.9 h. Lightcurve modulations are present that change slightly in phase and increase in amplitude from 0.015 to 0.07 mag over the three runs, suggestive of spot activity on either star. With  $r' \approx 19.98$  and  $r' - i' \approx 0.93$  we derive a primary mass of  $0.70M_{\odot}$ , spectral type K4V and distance 2.6 kpc, a possible cluster member. We also derive a secondary radius similar to the primary radius from our lightcurve modelling. Hence we classify this system as a grazing RS CVn type eclipsing binary consisting of a pair of K4V stars.

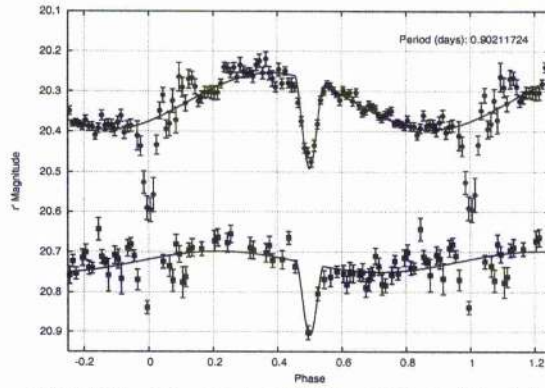
**Star 46271** shows no eclipses in the 1999-06 run. The 1999-07 and 2000-09 runs reveal a 0.902 d eclipsing binary (Figure 4.6(c)) with eclipses of different depths ( $\sim 0.04$  mag difference), where the primary eclipse is of depth 0.22 mag and duration 1.9 h. Lightcurve modu-



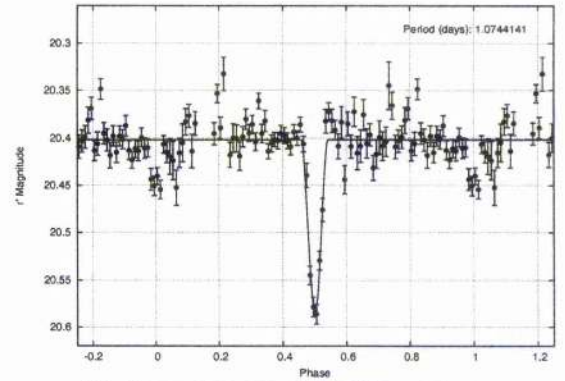
(a) 1262 - PL, BL & FTFIT - 2000-09 (Top), 1999-07 (Middle with +0.1 mag offset) & 1999-06 (Bottom with +0.2 mag offset)



(b) 22688 - PL, BL & FTFIT - 2000-09 (Top), 1999-07 (Middle with +0.2 mag offset) & 1999-06 (Bottom with +0.4 mag offset)

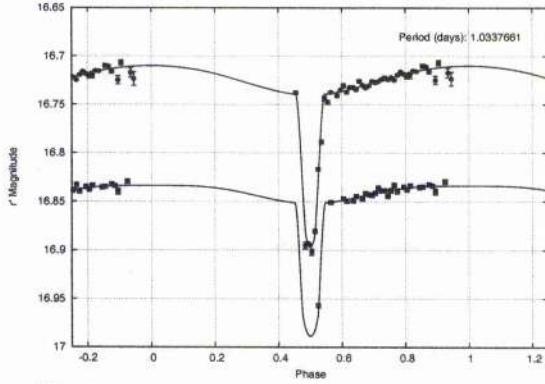


(c) 46271 - PL, BL & FTFIT - 2000-09 (Top) & 1999-07 (Bottom with +0.4 mag offset)

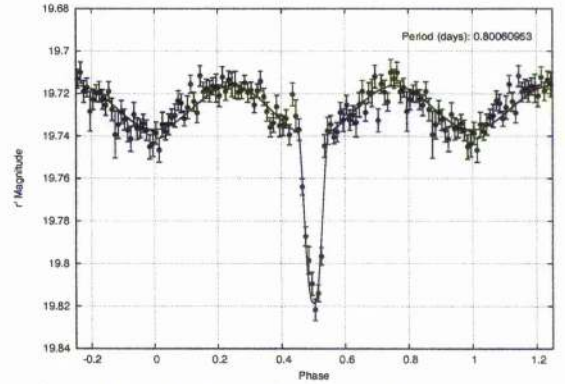


(d) 62983 - PL, BL & FTFIT - 2000-09

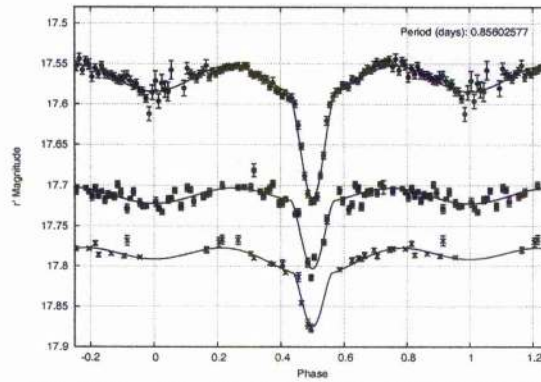
Figure 4.6: Eclipsing binaries exhibiting secondary eclipses.



(a) 5917 - PL, BL & FTFIT - 1999-07 (Top) & 1999-06 (Bottom with +0.1 mag offset)



(b) 21790 - PL, BL & FTFIT - 1999-06, 1999-07 & 2000-09



(c) 47171 - PL, BL & FTFIT - 2000-09 (Top), 1999-07 (Middle with +0.1 mag offset) & 1999-06 (Bottom with +0.2 mag offset)

Figure 4.7: Eclipsing binaries exhibiting ellipsoidal variations and heating effects.



lations are present that change in phase and increase in amplitude from 0.05 to 0.13 mag over the three runs, suggestive of spot activity on either star. With  $r' \approx 20.32$  and  $r' - i' \approx 1.03$  we derive a primary mass of  $0.66M_{\odot}$ , spectral type K5V and distance 2.5kpc, another possible cluster member. Again we derive a secondary radius similar to the primary radius from our lightcurve modelling. Hence we also classify this system as a grazing RS CVn type eclipsing binary consisting of a pair of K5V stars.

**Star 62983** only shows eclipses in the 2000-09 run due to its period of 1.07 d being close to an integer value. It is only on folding and binning the lightcurve data that the secondary eclipse becomes apparent (0.05 mag depth - Figure 4.6(d)). The primary eclipse is of depth 0.18 mag and duration 2.1 h. The star has  $r' \approx 20.40$  and  $r' - i' \approx 0.62$  suggesting a  $0.89M_{\odot}$  late G star that lies at  $d = 6.8\text{kpc}$ , twice the distance to the cluster. Our fits to the lightcurves suggest that the size of the secondary star is  $\sim 0.38R_{\odot}$ . We classify this system as an eclipsing binary, due to the absence of out-of-eclipse lightcurve variations, consisting of a late G star primary and an early M star secondary.

#### 4.2.5 Eclipsing Binaries Exhibiting Ellipsoidal Variations And Heating Effects

In this section we present 3 transit candidates that exhibit ellipsoidal variations and heating effects in their lightcurves which immediately implies that the companion is stellar. Figure 4.7 shows the folded lightcurve for each transit candidate along with the best fit transiting planet model using Model 3. The lightcurves from different runs are offset vertically (in magnitude) from each other in order to highlight any changes in the amplitude of the out-of-eclipse variations.

**Star 5917** only shows eclipses in the 1999-06 and 1999-07 runs due to its period of 1.03 d being close to an integer value. On folding and binning the lightcurve data (Figure 4.7(a)), it becomes apparent that the phase coverage is not sufficient to detect any secondary eclipse that might be present. The primary eclipse is of depth 0.16 mag and duration 2.2 h, and the lightcurve modulations seem to increase in amplitude from 0.01 to 0.02 mag over the space of 1 month. The star is the brightest transit candidate with  $r' \approx 16.72$  and  $r' - i' \approx 0.55$

suggesting a  $0.94M_{\odot}$  early G star that lies at  $d = 1.6\text{kpc}$ , in front of the cluster. We find that the companion has a radius of  $0.33R_{\odot}$  from our fits to the lightcurve data. Hence we classify this system as an Algol type eclipsing binary consisting of an early G star primary and a M3V star secondary.

**Star 21790** exhibits 0.08 mag eclipses of duration 1.8 h in all three runs (Figure 4.7(b)). The lightcurve data folded and binned on a period of 0.801 d clearly shows out-of-eclipse ellipsoidal variations of amplitude 0.02 mag. Evidence for the heating effect is negligible (Table 4.3) and no secondary eclipse seems to be present. With  $r' \approx 19.73$  and  $r' - i' \approx 0.56$  we derive a primary  $0.98M_{\odot}$  G2V star like our Sun that lies at  $d = 6.6\text{kpc}$ , twice the distance to the cluster. Our lightcurve modelling reveals that the companion has a radius of  $0.26R_{\odot}$ . Hence we classify this system as an Algol type eclipsing binary consisting of a Sun-like star orbited by a M4V star.

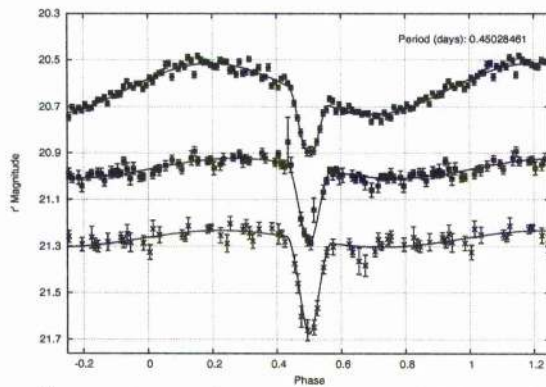
**Star 47171** shows eclipses that increase in depth from 0.07 to 0.13 mag and lightcurve modulations that increase slightly in amplitude from 0.04 to 0.05 mag over the three runs (Figure 4.7(c)). The period is 0.856 d with an eclipse duration of 3.0 h. The star has  $r' \approx 17.57$  and  $r' - i' \approx 0.43$  from which we derive a primary mass of  $1.33M_{\odot}$ . We also derive a companion radius of  $0.58R_{\odot}$  from the fits to the lightcurve data. Hence we classify this system as an Algol type eclipsing binary consisting of a F star primary and a K7V star secondary that lies at  $d = 5.6\text{kpc}$ , beyond the cluster.

#### 4.2.6 A Possible Long Period Cataclysmic Variable

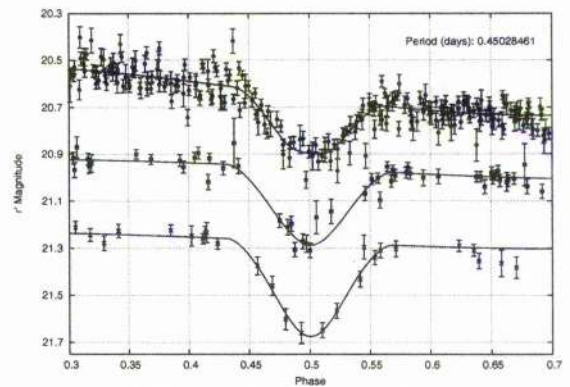
**Star 711** is a 10.8 h eclipsing binary (Figure 4.8) that has round-bottomed eclipses lasting 0.1 in phase and orbital modulations that peak near phase 0.2. With  $r' \approx 20.62$  mag and  $r' - i' \approx 1.13$  mag, the star falls close to the cluster main sequence (Figure 3.4(a)). The colour index is consistent with a  $0.62M_{\odot}$  K7V star at  $d = 2.4\text{kpc}$ , a possible cluster member. Over the three runs the orbital modulations increase in amplitude from 0.1 mag to 0.2 mag, while the eclipse depth decreases from 0.40 mag to 0.24 mag.

The orbital modulation could arise from spots on one or both stars, though this would require a preferred longitude that remains stable over 15 months. The orbital phasing is





(a) 711 - PL, BL & FTFIT - 2000-09 (Top), 1999-07 (Middle with +0.3 mag offset) & 1999-06 (Bottom with +0.6 mag offset)



(b) 711 - CPL & FTFIT - 2000-09 (Top), 1999-07 (Middle with +0.3 mag offset) & 1999-06 (Bottom with +0.6 mag offset)

Figure 4.8: Possible long period cataclysmic variable.

consistent with that of an “orbital hump” that is often seen in quiescent dwarf novae, arising from the anisotropic emission of a “hotspot” on the rim of an accretion disk where the mass transfer stream from the companion star feeds material into the disk. The relatively shallow eclipse would then imply a moderate inclination so that the donor star eclipses only the near rim of the disk and possibly the hotspot. The eclipse shape is more symmetric than would be expected for eclipses of a hotspot, however, and a hotspot eclipse would become deeper rather than shallower as the orbital modulation increased. We are unable to decide which interpretation may be correct and recommend follow-up observations to resolve this ambiguity. In any case the eclipse is too deep to be attributed to a planetary transit.

#### 4.2.7 More Eclipsing Binaries

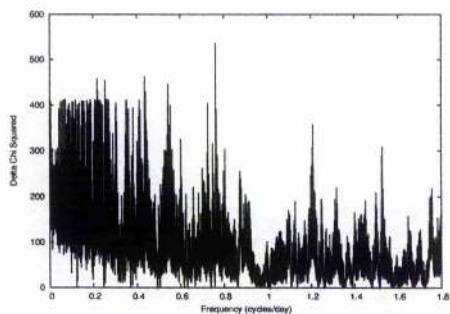
In this section we present 6 transit candidates exhibiting neither easily discernible secondary eclipses nor orbital modulations consistent with ellipsoidal variations and heating effects. Figures 4.9 to 4.11 show for each star an eclipse periodogram, a chi squared contour map, a folded and binned lightcurve with the best fit transiting planet model, and an unbinned close-up of the folded lightcurve around the primary eclipse along with the best fit model. For star 7628 the lightcurves from different runs are offset vertically to highlight changes

in the amplitude and phase of the out-of-eclipse variations. We rule out the transiting planet model for stars 6995, 7628, 22738, 50313 and 73852 because the full transit fit yields a companion radius greater than  $0.2R_{\odot}$ . For star 64804 the full transit fit admits  $R_c < 0.2R_{\odot}$  but only for periods  $P < 1.1$  d that are ruled out and hence the transiting planet model is ruled out also for star 64804.

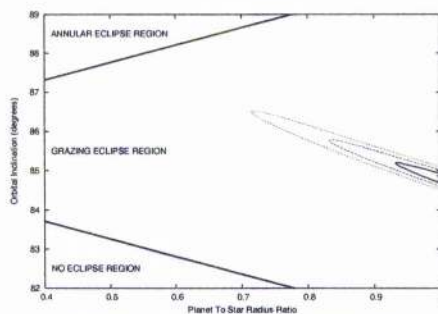
**Star 6995** shows 4 V-shaped eclipses of depth 0.39 mag and duration 1.3 h over the three runs. Figures 4.9(c) and 4.9(d) show the lightcurve folded on the 1.31 d period as derived from the eclipse periodogram (Figure 4.9(a)). With  $r' \approx 21.12$  mag and  $r' - i' \approx 1.85$  mag we derive a primary  $0.24M_{\odot}$  M5V star at  $d \sim 750$  pc, in front of the cluster. The full transit fit (Figure 4.9(b)) reveals that the companion is the same size as the primary and that the eclipses are grazing. Hence the period is actually 2.62 d and we classify the system as a grazing eclipsing binary consisting of a pair of M5V stars, an interesting discovery in that few such systems are known.

**Star 7628** has a period of 1.76 d and eclipses of depth 0.16 mag and duration 1.7 h (Figures 4.9(e), 4.9(g) & 4.9(h)). The sinusoidal out-of-eclipse variations change in phase and decrease in amplitude from 0.07 to 0.04 mag over the three runs, suggestive of spot activity on the primary star. With  $r' \approx 19.42$  mag and  $r' - i' \approx 1.30$  mag we derive a primary mass of  $0.54M_{\odot}$ , spectral type K9V and distance 1.0 kpc, placing the system in front of the cluster. We derive a companion radius of  $0.33R_{\odot}$  and inclination  $83.7^{\circ}$  from the full transit fit suggesting that the eclipses are grazing (Figures 4.9(f)). We classify this system as a grazing RS CVn type eclipsing binary consisting of a K9V primary and a M star secondary.

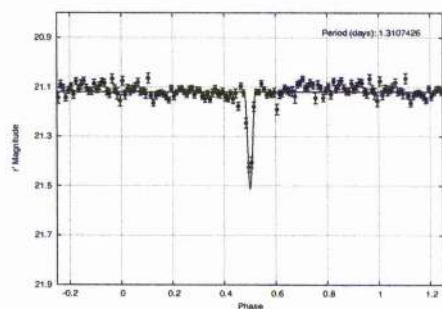
**Star 22738** shows three partially observed eclipses, two of them in the 2000-09 run. The eclipses have a depth of 0.17 mag and duration 2.9 h with a period of 2.77 d (Figures 4.10(a), 4.10(c) & 4.10(d)). The 2000-09 lightcurve exhibits modulations of amplitude 0.02 mag that are out of phase with the eclipses, suggestive of star spots on the primary star. In fact, the folded and binned lightcurve (Figure 4.10(c)) reveals a possible 0.01 mag secondary eclipse. The star has  $r' \approx 17.65$  mag and  $r' - i' \approx 0.64$  mag suggesting a  $0.81M_{\odot}$  G9V primary star that lies at  $d = 1.6$  kpc, in front of the cluster. The full transit fit (Figure 4.10(b)) reveals that the eclipses are annular, nearly central ( $i \approx 89.9^{\circ}$ ) and that



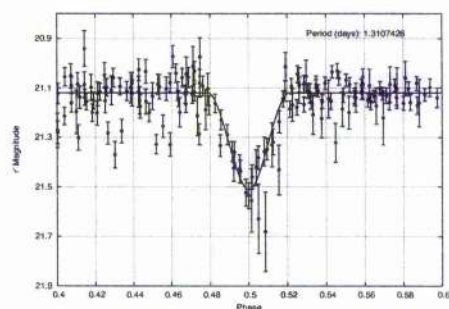
(a) 6995 - EP - 1999-06, 1999-07 & 2000-09



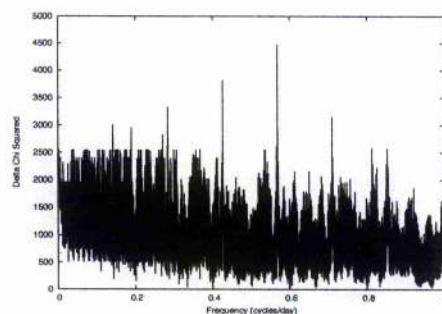
(b) 6995 - CM - 1999-06, 1999-07 & 2000-09



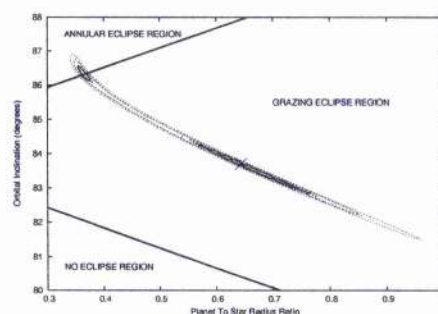
(c) 6995 - PL, BL & FTFIT - 1999-06, 1999-07 & 2000-09



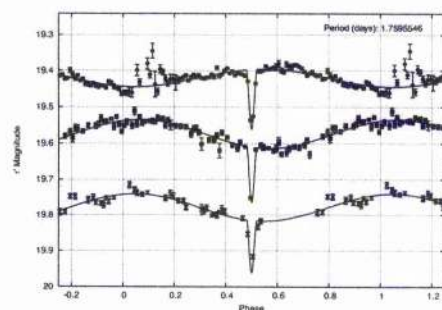
(d) 6995 - CPL & FTFIT - 1999-06, 1999-07 & 2000-09



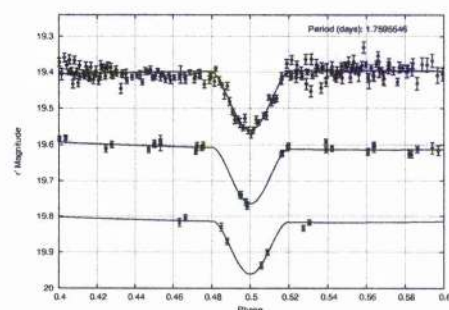
(e) 7628 - EP - 1999-06, 1999-07 & 2000-09



(f) 7628 - CM - 2000-09



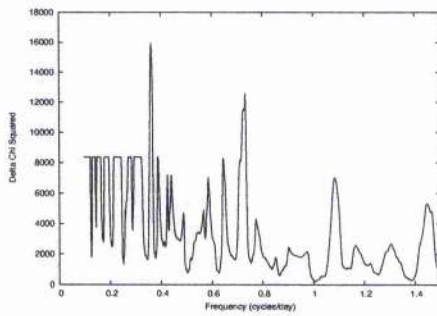
(g) 7628 - PL, BL & FTFIT - 2000-09 (Top), 1999-07 (Middle with +0.2 mag offset) & 1999-06 (Bottom with +0.4 mag offset)



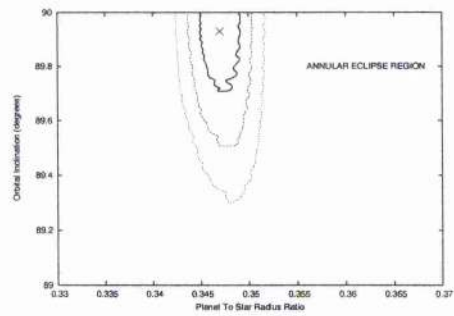
(h) 7628 - CPL & FTFIT - 2000-09 (Top), 1999-07 (Middle with +0.2 mag offset) & 1999-06 (Bottom with +0.4 mag offset)

Figure 4.9: Stars 6995 and 7628.

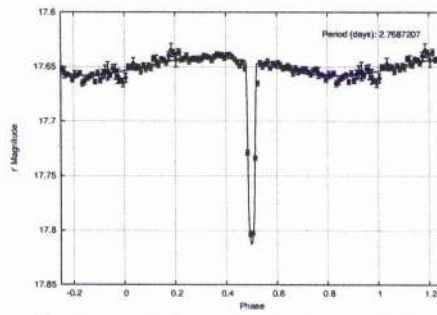




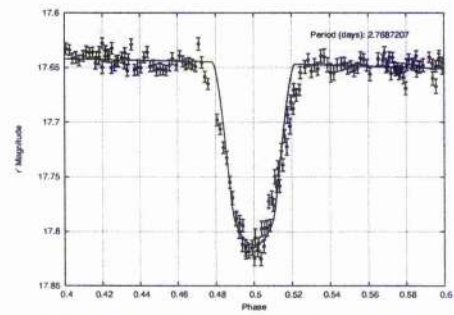
(a) 22738 - EP - 2000-09



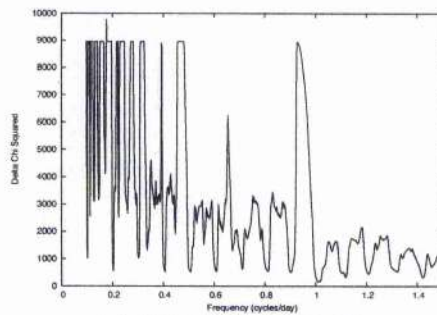
(b) 22738 - CM - 2000-09



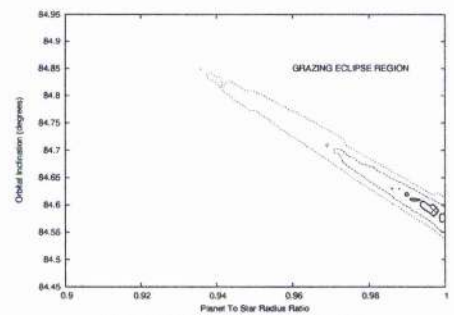
(c) 22738 - PL, BL &amp; FTFIT - 2000-09



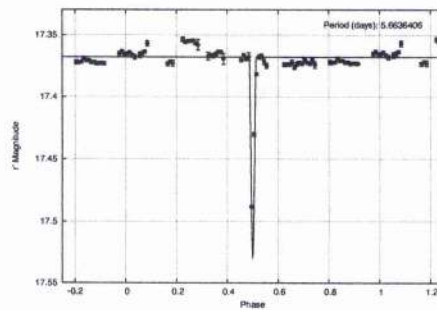
(d) 22738 - CPL &amp; FTFIT - 2000-09



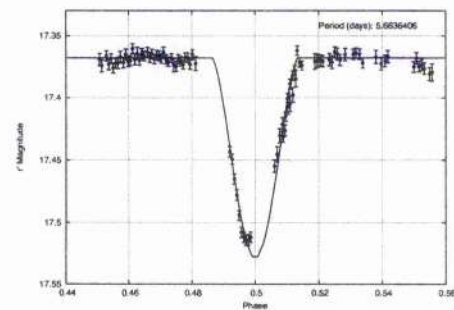
(e) 50313 - EP - 2000-09



(f) 50313 - CM - 2000-09



(g) 50313 - PL, BL &amp; FTFIT - 2000-09



(h) 50313 - CPL &amp; FTFIT - 2000-09

Figure 4.10: Stars 22738 and 50313.

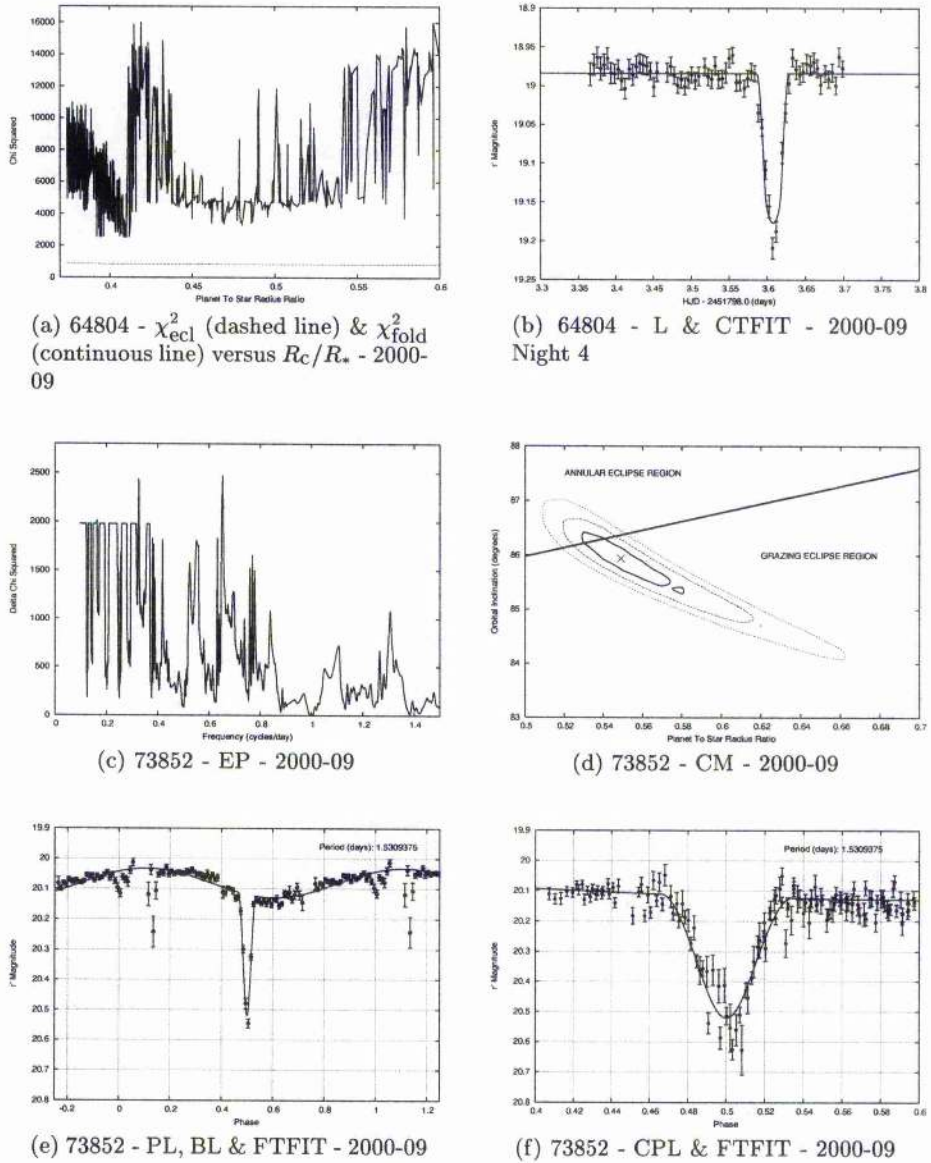


Figure 4.11: Stars 64804 and 73852.

the companion has a radius of  $0.27R_{\odot}$ . We classify this system as an annular RS CVn type eclipsing binary consisting of a G9V primary and a M star secondary.

**Star 50313** shows four partially observed eclipses, two of them in the 2000-09 run. The eclipses have a depth of 0.16 mag and duration 3.7 h with a period of 5.66 d (Figures 4.10(e), 4.10(g) & 4.10(h)). The 2000-09 lightcurve exhibits erratic out-of-eclipse variations of amplitude  $\sim 0.02$  mag which we were unable to model. Such variations are most likely due to stellar activity/star spots on the primary star. With  $r' \approx 17.37$  mag and  $r' - i' \approx 0.62$  mag we derive a primary  $0.82M_{\odot}$  G9V star that lies at  $d = 1.5$  kpc, in front of the cluster. The full transit fit (Figure 4.10(f)) reveals that the companion is the same size as the primary and that the eclipses are grazing. Hence the period is actually 11.33 d and we classify the system as a grazing RS CVn eclipsing binary consisting of a pair of G9V stars.

**Star 73852** shows primary eclipses of depth 0.40 mag and duration 2.4 h, lightcurve modulations of amplitude 0.09 mag that are out of phase with the eclipses (suggestive of star spots on the primary) and a possible secondary eclipse of depth 0.05 mag (Figures 4.11(e) & 4.11(f)). The eclipse periodogram gives a period of 1.53 d (Figure 4.11(c)). The star has  $r' \approx 20.08$  mag and  $r' - i' \approx 0.80$  mag giving a primary star mass of  $0.74M_{\odot}$  and  $d = 3.4$  kpc, suggesting that it is a K2V star beyond the cluster. We derive a companion radius of  $0.39R_{\odot}$  and inclination  $85.9^{\circ}$  from the full transit fit suggesting that the eclipses are just grazing (Figure 4.11(d)). We classify this system as a grazing RS CVn type eclipsing binary consisting of a K2V primary and an early M star secondary.

**Star 64804** is a difficult case in that the lightcurve data show one eclipse in the 1999-06 run with only 2 data points during the eclipse, and one well sampled eclipse in the 2000-09 run (Figure 4.11(b)). The eclipse is V-shaped of depth 0.2 mag and duration 1.0 h suggesting that it is likely to be a grazing eclipse. With  $r' \approx 18.98$  mag and  $r' - i' \approx 1.52$  mag we find that the primary is a  $0.40M_{\odot}$  M2V star that lies at only  $\sim 530$  pc. A central transit fit to the 2000-09 data (Figure 4.11(b)) yields a minimum companion radius of  $0.141R_{\odot}$  due to the small size of the primary star ( $0.38R_{\odot}$ ) from which we cannot rule out the transiting planet model. Analysis of the shape of the single eclipse in the 2000-09 data is possible due to the good time sampling of the observations and such an analysis may reveal whether the eclipse is the result of an annular occultation by a smaller companion or a grazing

occultation by a larger companion. Also, we may attempt to predict the orbital period  $P$  of the planetary companion as a function of the impact parameter  $b = a \cos i / R_*$  of the eclipse and subsequently use the lightcurve data from the whole run to determine which periods, and hence which values of  $b$ , may be ruled out.

For star 64804 we made a grid for the impact parameter  $b$  from 0.0 to 1.0. For each value of  $b$  we fitted Model 1 to the single eclipse in the 2000-09 run, using the lightcurve data from the whole run, in order to determine a time of mid-transit  $t_0$ , a constant magnitude  $m_0$ , a planetary radius  $R_c$  and a transit duration  $\Delta t$ . The chi squared  $\chi_{\text{ecl}}^2$  of the fit was also calculated. The duration of a transit event (see Equation 2.19) is given by:

$$\Delta t = \frac{PR_*}{\pi a} \sqrt{\left(1 + \frac{R_c}{R_*}\right)^2 - b^2} \quad (4.4)$$

We already know  $M_*$  and  $R_*$ , and since we have  $R_c$  and  $\Delta t$  as functions of  $b$  from our fits of the single eclipse, we may use Equations 1.8 and 4.4 to estimate  $P$  (or  $a$ ) as a function of  $b$  for the transiting planet model. For each value of  $b$  we folded the 2000-09 lightcurve of star 64804 on the predicted period  $P$  using the fitted  $t_0$ , and calculated a new chi squared  $\chi_{\text{fold}}^2$  using the fit to the single eclipse with period  $P$ . In general, if the predicted period is such that none of the folded lightcurve data falls during the eclipse, then  $\chi_{\text{fold}}^2 = \chi_{\text{ecl}}^2$ . However, if the predicted period is such that some of the folded lightcurve data does fall during the eclipse, then  $\chi_{\text{fold}}^2 \gg \chi_{\text{ecl}}^2$ , ruling out that particular period, impact parameter and corresponding eclipse solution.

Figure 4.11(a) shows a plot of  $\chi_{\text{ecl}}^2$  versus  $R_c/R_*$  (dashed line), which appears constant due to the scale on the y-axis. The continuous line is a plot of  $\chi_{\text{fold}}^2$  versus  $R_c/R_*$ , which clearly shows that  $\chi_{\text{fold}}^2 \gg \chi_{\text{ecl}}^2$  for all values of  $R_c$  from  $0.141R_\odot$  (the minimum companion radius with  $b = 0$  and  $R_c/R_* = 0.373$ ) to  $0.227R_\odot$  ( $b = 0.92$  and  $R_c/R_* = 0.600$ ). This is due to the fact that the predicted period is less than 1.10 d for these values of  $b$ . This demonstrates that the transiting planet model is inconsistent with our observational data for this star, and hence a stellar companion is favoured. Table 4.4 presents the results of the central transit fit to the eclipse during the 2000-09 run. We have classified this system as an eclipsing binary.



#### 4.2.8 INT-7789-TR-1 (Star 45134)

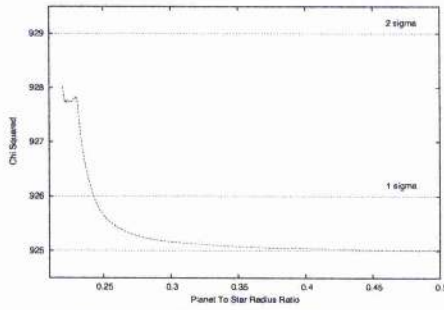
**Star 45134** exhibits one poorly sampled partially observed eclipse during the 1999-07 run and one well sampled fully observed eclipse of depth 0.07 mag and duration 6.0 h in the 2000-09 run (Figure 4.12(b)), and as a result we were unable to determine a period for the system. The star has  $r' \approx 20.70$  mag and  $r' - i' \approx 0.63$  mag from which we derive a  $0.87M_{\odot}$  late G star primary that lies far behind the cluster ( $d = 7.3\text{kpc}$ ). A central transit fit to the 2000-09 data yields a minimum companion radius of  $0.188R_{\odot}$  from which we cannot rule out the transiting planet model.

Analysing the single eclipse in the 2000-09 run using the same method as for star 64804 in Section 4.2.7 yields a predicted period of 7.0 d for  $b = 0$  that increases rapidly with increasing values of  $b$ . In fact  $\chi_{\text{fold}}^2 = \chi_{\text{ecl}}^2$  for all  $b > 0.09$ . Figure 4.12(a) shows a plot of  $\chi_{\text{ecl}}^2$  versus  $R_c/R_*$  (dashed line) for  $b = 0$  ( $R_c/R_* = 0.223$ ) to  $b = 1.13$  ( $R_c/R_* = 0.500$ ). The minimum value of  $\chi_{\text{ecl}}^2$  obtained is  $\chi_{\text{ecl}}^2 = 925.0$  corresponding to  $R_c/R_* = 0.500$ , and this is marked on Figure 4.12(a) as a horizontal shorter dashed line, along with the chi squared values  $\chi_{\text{ecl}}^2 = 925.0 + 1.0$  and  $\chi_{\text{ecl}}^2 = 925.0 + 4.0$  corresponding to the 1 and  $2\sigma$  confidence levels.

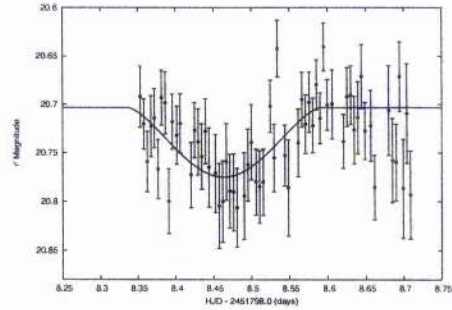
One can see from Figure 4.12(a) that  $R_c/R_* \geq 0.243$  with a  $1\sigma$  confidence. This is equivalent to stating that  $R_c \geq 0.205R_{\odot}$  with a  $1\sigma$  confidence. As a result, we can only rule out the transiting planet model for this transit candidate at the  $1\sigma$  level and therefore further observations are required to confirm the conclusion that this system is an eclipsing binary. Table 4.4 presents the solution corresponding to the minimum value of  $\chi_{\text{ecl}}^2$  including the predicted period and inclination, and Figure 4.12(b) shows a plot of this solution along with the lightcurve data for the night on which the eclipse occurs.

#### 4.2.9 INT-7789-TR-2 (Star 46691)

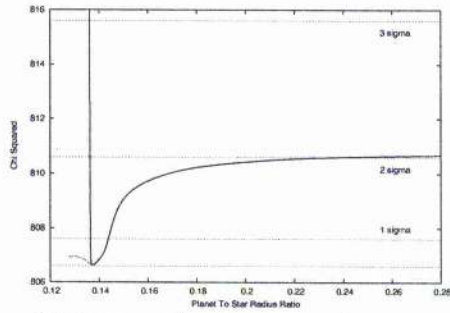
**Star 46691** shows a single 0.02 mag eclipse of duration 2.5 h during the 2000-09 run (Figure 4.12(d)). With  $r' \approx 18.02$  mag and  $r' - i' \approx 0.47$  mag we find that the primary is a  $1.20M_{\odot}$  F star at  $d = 5.3\text{kpc}$ , behind the cluster. A central transit fit to the 2000-09 data yields a minimum companion radius of  $0.174R_{\odot}$  from which we cannot rule out the



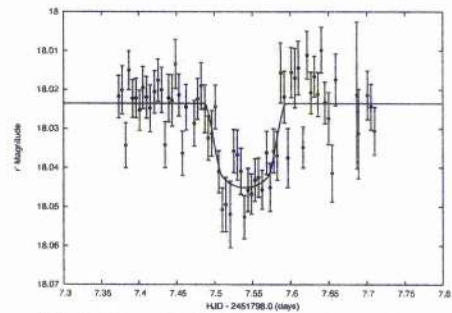
(a) 45134 -  $\chi^2_{\text{ecl}}$  (dashed line) versus  $R_c/R_*$  - 2000-09



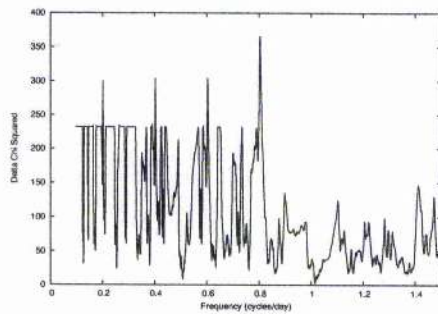
(b) 45134 - L & Single Eclipse Fit - 2000-09 Night 9



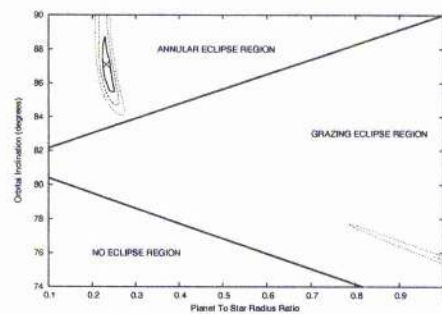
(c) 46691 -  $\chi^2_{\text{ecl}}$  (dashed line) versus  $R_c/R_*$  &  $\chi^2_{\text{fold}}$  (continuous line) versus  $R_c/R_*$  - 2000-09



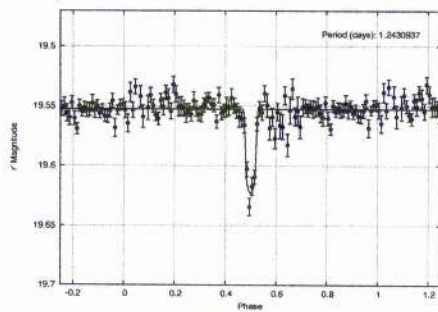
(d) 46691 - L & Single Eclipse Fit - 2000-09 Night 8



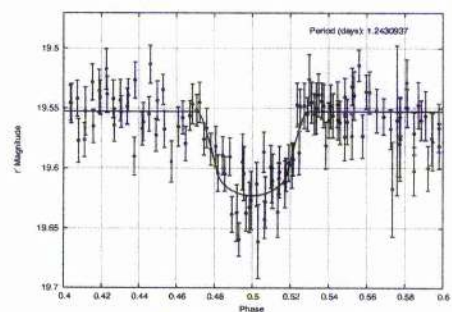
(e) 49512 - EP - 2000-09



(f) 49512 - CM - 2000-09

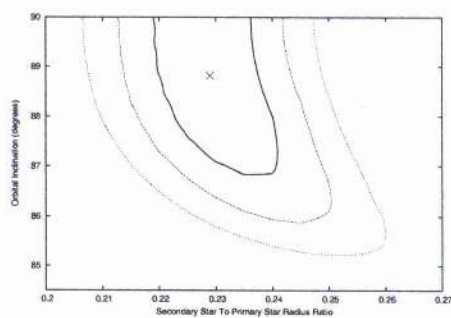


(g) 49512 - PL, BL & FTFIT - 2000-09

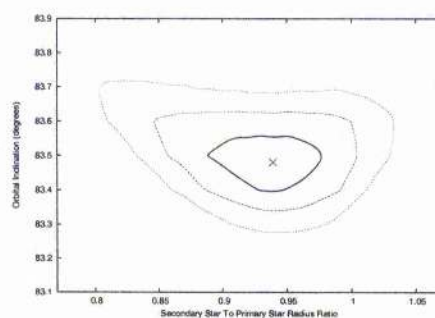


(h) 49512 - CPL & FTFIT - 2000-09

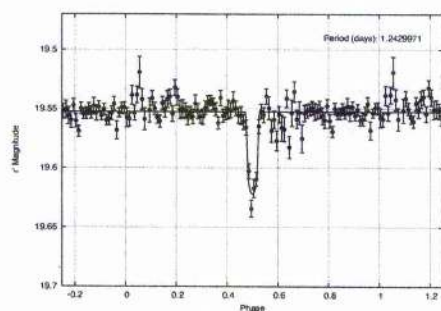
Figure 4.12: Planetary transit candidates 45134, 46691 and 49512.



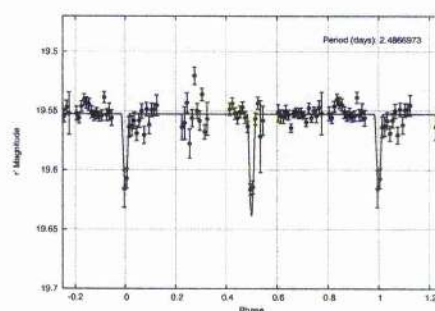
(a) 49512 - CM - 2000-09



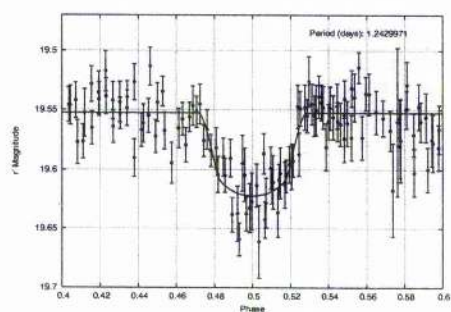
(b) 49512 - CM - 2000-09



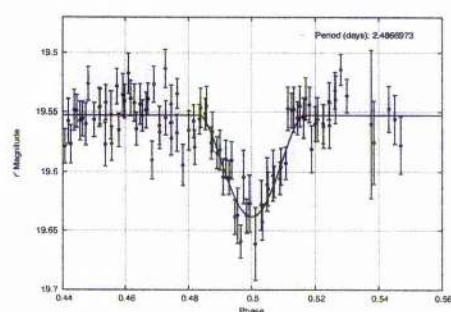
(c) 49512 - PL, BL &amp; Eclipsing Binary Fit - 2000-09



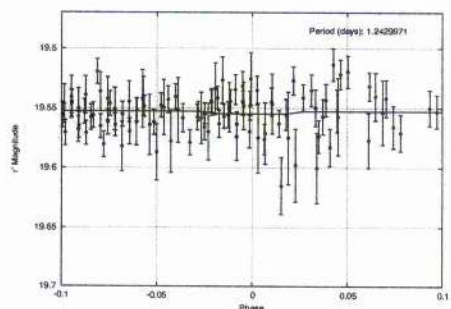
(d) 49512 - PL, BL &amp; Eclipsing Binary Fit - 2000-09



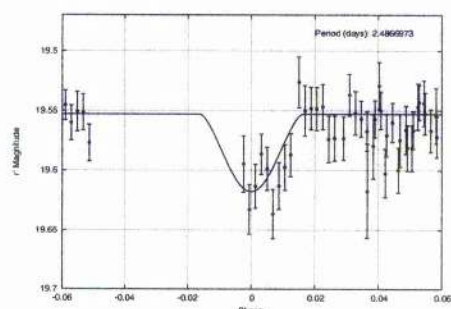
(e) 49512 - CPL &amp; Eclipsing Binary Fit - 2000-09



(f) 49512 - CPL &amp; Eclipsing Binary Fit - 2000-09



(g) 49512 - CPL &amp; Eclipsing Binary Fit - 2000-09



(h) 49512 - CPL &amp; Eclipsing Binary Fit - 2000-09

Figure 4.13: Eclipsing binary fits for star 49512.

transiting planet model.

Table 4.5: Star, companion and lightcurve properties for INT-7789-TR-3 (star 49512) as obtained from the various fits detailed in Section 4.2.10. Column 4 is calibrated  $r'$  magnitude.  $M_c$  and  $R_c$  are the companion mass and radius respectively. The ratio  $f_c/f_*$  is the flux ratio of the companion to the primary star in the Sloan  $r'$  waveband. The quantity  $d$  is the distance to the system.

Model	$t_0 - 2451000.0$ (d)	$P$ (d)	$m_0$ ( $r'$ mag)	$\Delta m$ (mag)	$\Delta t$ (h)	$i$ ( $^\circ$ )	$M_*$ ( $M_\odot$ )	$R_*$ ( $R_\odot$ )	$M_c$ ( $M_\odot$ )	$R_c$ ( $R_\odot$ )	$f_c/f_*$ (Sloan $r'$ )	$d$ (pc)	$\chi^2$
A	800.6201(7)	2.4867(116)	19.553	0.086	1.93	83.48(6)	0.701(5)	0.670(5)	0.661(30)	0.629(29)	0.669	2788(6)	725.88
B	800.6203(7)	1.2431(58)	19.553	0.070	1.72	87.1(1.2)	0.679(4)	0.649(4)	0.000	0.151(7)	0.000	1995(5)	729.99
C	800.6204(7)	1.2430(58)	19.553	0.070	1.68	88.8(1.5)	0.682(4)	0.651(4)	0.126(6)	0.149(6)	0.0019	2021(5)	731.73

A = Similar Size Stellar Companion. B = Planetary Companion. C = Small Stellar Companion.

Applying the same analysis as for star 64804 in Section 4.2.7 to the single eclipse yields a predicted period of 0.62 d for  $b = 0$  that increases slowly with increasing values of  $b$ . For  $b < 0.66$ ,  $\chi_{\text{fold}}^2 \gg \chi_{\text{ecl}}^2$  and for  $b \geq 0.66$ ,  $\chi_{\text{fold}}^2$  oscillates between the states  $\chi_{\text{fold}}^2 \gg \chi_{\text{ecl}}^2$  and  $\chi_{\text{fold}}^2 = \chi_{\text{ecl}}^2$ . Hence we can be sure that  $b \geq 0.66$ , which corresponds to  $R_c/R_* \geq 0.137$ . In Figure 4.12(c) we plot  $\chi_{\text{ecl}}^2$  versus  $R_c/R_*$  (dashed line) and  $\chi_{\text{fold}}^{2'}$  versus  $R_c/R_*$  (continuous line) where:

$$\chi_{\text{fold}}^{2'} = \begin{cases} \chi_{\text{fold}}^2 & \text{if } b < 0.66 \\ \chi_{\text{ecl}}^2 & \text{if } b \geq 0.66 \end{cases} \quad (4.5)$$

The minimum value of  $\chi_{\text{ecl}}^2$  obtained is  $\chi_{\text{ecl}}^2 = 806.6$  corresponding to  $R_c/R_* = 0.137$ , and this is marked on Figure 4.12(c) as a horizontal shorter dashed line, along with the chi squared values  $\chi_{\text{ecl}}^2 = 806.6 + 1.0$ ,  $\chi_{\text{ecl}}^2 = 806.6 + 4.0$  and  $\chi_{\text{ecl}}^2 = 806.6 + 9.0$  corresponding to the 1, 2 and 3 $\sigma$  confidence levels.

One can see from Figure 4.12(c) that  $0.137 \leq R_c/R_* \leq 0.144$  with a 1 $\sigma$  confidence. This is equivalent to stating that  $R_c = 0.185_{-0.000}^{+0.009} R_\odot$ . Hence the conclusion at the 1 $\sigma$  level is that this is a possible transiting planet in orbit around a  $1.20 M_\odot$  F star that merits follow-up observations. Table 4.4 presents the solution corresponding to the minimum value of  $\chi_{\text{ecl}}^2$  including the predicted period and inclination, and Figure 4.12(d) shows a plot of this solution along with the lightcurve data for the night on which the eclipse occurs.

#### 4.2.10 INT-7789-TR-3 (Star 49512)

**Star 49512** exhibits two fully observed eclipses and one partially observed eclipse during the 2000-09 run. The eclipses have a depth of 0.07 mag and duration 1.7 h with a period of 1.24 d (Figures 4.12(e), 4.12(g) & 4.12(h)). The star has  $r' \approx 19.55$  mag and  $r' - i' \approx 0.97$  mag from which we derive a  $0.68 M_\odot$  primary star of spectral type K5V that lies slightly in front of the cluster ( $d = 2.0$  kpc). A full transit fit to the 2000-09 lightcurve data yields a best fit companion radius of  $0.151 \pm 0.007 R_\odot$  consistent with the radius of a transiting planet (Figure 4.12(f)). This solution is reported in Table 4.4, and the  $\chi^2$  of the fit is 729.99.

However, there are two other models for star 49512 that should be considered to see



if they produce a better  $\chi^2$  value for the fit to the lightcurve data. It is possible that the companion is a smaller and less luminous star than the primary star that produces a secondary eclipse which is not visible in the lightcurve folded at the  $\sim 1.24$  day period, and it is also possible that the companion is a star of similar size and luminosity to the primary star and that the system actually has an orbital period of  $\sim 2.48$  days. In order to test these models for star 49512 we have developed an eclipsing binary model based on the same assumptions as for the star and planet system presented in Section 4.2.1 except that we assume that the companion is now luminous and massive, and that our theoretical main sequence relationships adopted in Section 3.4.3 apply to the companion. The model is represented mathematically by:

$$f_4(t) = f_0(1.0 - f_c(t) - f_*(t)) \quad \text{Model 4} \quad (4.6)$$

The function  $f_4(t)$  is the predicted stellar flux at time  $t$ ,  $f_0$  is a constant flux value,  $f_c(t)$  is the fraction of the total stellar flux obscured by the companion at time  $t$  and  $f_*(t)$  is the fraction of the total stellar flux obscured by the primary star at time  $t$ . The function  $f_4(t)$  is calculated in a numerical fashion by creating a grid for the observed stellar disks and calculating the flux from each grid element taking into account the apparent position of the companion at time  $t$ , the different surface brightnesses of each star and the effect of linear limb darkening with  $u = 0.5$ .

The eclipsing binary model has five parameters to optimise: orbital period  $P$ , time of mid-eclipse  $t_0$ , orbital inclination  $i$ , companion to primary star radius ratio  $R_c/R_*$  and a constant magnitude  $m_0$ . We fitted this model to the lightcurve of star 49512 by calculating the  $\chi^2$  for a grid in  $i$  and  $R_c/R_*$ , in the same way as for the transiting planet model. The added subtlety is that for each value of  $R_c/R_*$ , we had to recalculate the distance  $d$  to the system, the values of  $M_*$  and  $R_*$  for the primary star, and the mass of the companion  $M_c$ . This was done by constructing a theoretical binary main sequence for the current value of  $R_c/R_*$  and then finding the distance  $d$  such that this model passes through the position of star 49512 in the colour-magnitude domain. The initial value of  $P$  was either 1.24 d or 2.48 d corresponding to the small or similar size stellar companion models respectively.



Table 4.5 reports the results of these fits including the results of the fit of the transiting planet model for comparison. Figure 4.13 shows a chi squared contour map, a folded and binned lightcurve with the best fit eclipsing binary model, an unbinned close-up of the folded lightcurve around the primary eclipse along with the best fit model and another unbinned close-up around the secondary eclipse along with the best fit model. The left hand column of diagrams in Figure 4.13 applies to the case of the small stellar companion and the right hand column of diagrams in Figure 4.13 applies to the case of the similar size stellar companion.

Table 4.5 shows that the best model for star 49512 is the eclipsing binary model with a similar size stellar companion since this model attains the smallest  $\chi^2$  of 725.88. All three models require exactly 5 parameters to be optimised and hence we calculate a likelihood ratio of  $\sim 7.8$  for the eclipsing binary model with a similar size stellar companion compared to the transiting planet model, and we calculate a likelihood ratio of  $\sim 18.6$  for the eclipsing binary model with a similar size stellar companion compared to the eclipsing binary model with a small stellar companion. Finally, we calculate a likelihood ratio of  $\sim 2.4$  for the transiting planet model compared to the eclipsing binary model with a small stellar companion. Hence our conclusion is that this system is most likely to be a grazing eclipsing binary with period 2.49 d consisting of a K4V star primary and a K5V star secondary that lies at  $d = 2.8\text{kpc}$ , slightly behind the cluster. However, further observations will be required confirm this conclusion and categorically rule out the transiting planet model.

#### 4.2.11 Finding Charts

In order to help facilitate follow-up observations of the eclipsing binaries and transit candidates presented in the preceding sections, we supply finding charts in Figure 4.14. Each stamp is a  $27'' \times 27''$  section of the relevant reference frame where North is up and East is to the right. Each eclipsing binary/transit candidate lies at the centre of its stamp and is marked by a cross.

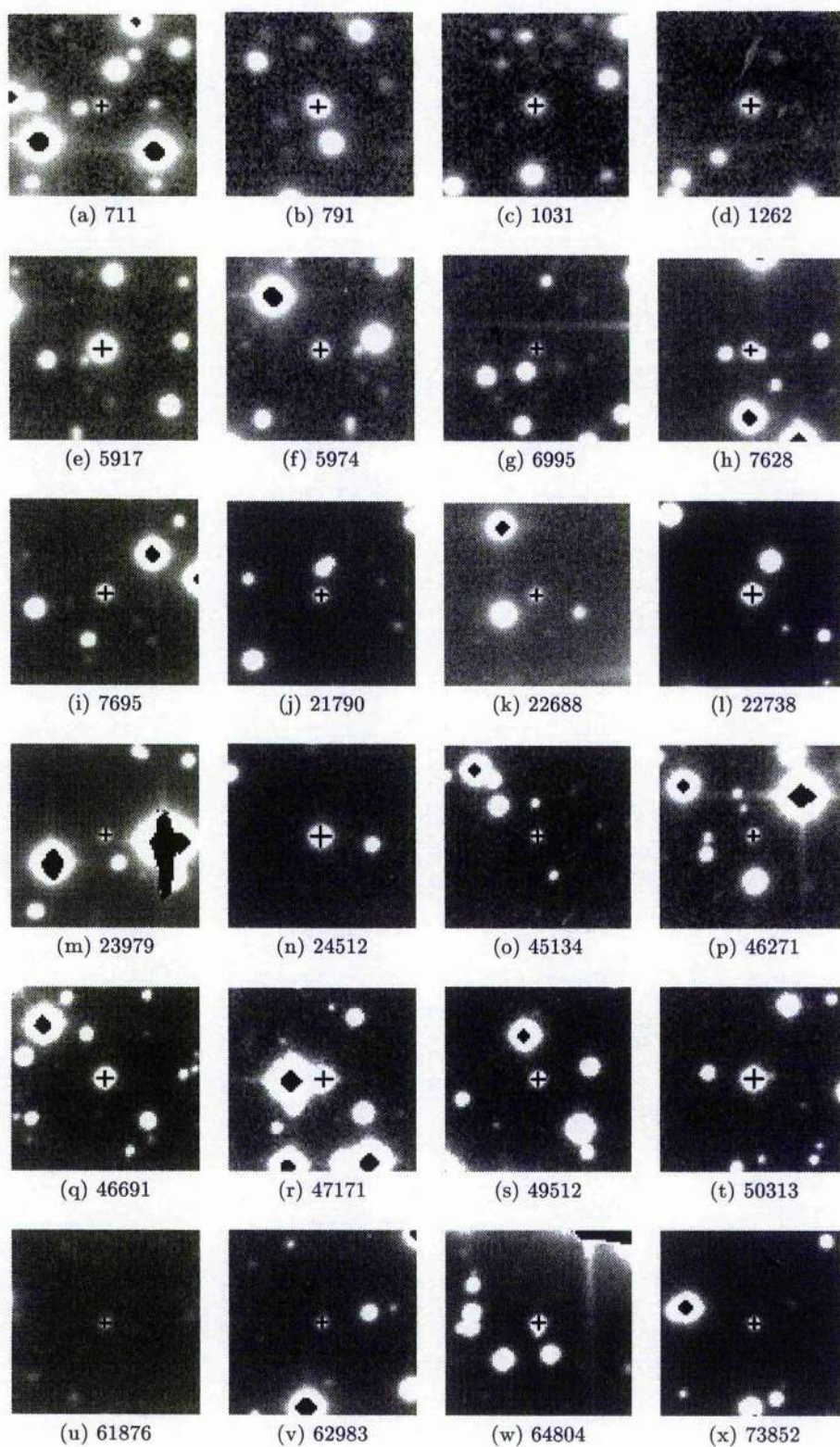


Figure 4.14: Finding charts taken from the reference frames. North is up and East is to the right. The stamps are of size  $27'' \times 27''$ . Each eclipsing binary/transit candidate lies at the centre of its stamp and is marked by a cross.

### 4.3 Summary

In this chapter we have introduced our transit detection algorithm, based on a matched filter algorithm, which was used to identify 2182 raw transit candidates from the lightcurves of the  $\sim 33000$  stars in our sample. We have described the data quality tests that we used to reject false transit candidates which left us with 24 remaining transit candidate lightcurves. The theoretical transiting planet model has been presented, along with the procedure used for modelling the lightcurves of the transit candidates.

We have been able to rule out the transiting planet model for 21 of the transit candidates using various robust arguments. For 2 candidates, INT-7789-TR-1 and INT-7789-TR-3, we have been unable to decide on their nature, although it seems most likely that they are eclipsing binaries as well. We have presented one candidate, INT-7789-TR-2, exhibiting a single eclipse for which we derive a radius of  $1.81^{+0.09}_{-0.00} R_J$ . Three candidates remain that require follow-up observations in order to determine their nature. Finally we have plotted finding charts in order to aid any follow-up observations of the eclipsing binaries and transit candidates.

# 5

---

## Limits On The Hot Jupiter Fraction In The Field Of NGC 7789

### 5.1 Introduction

In the previous Chapter we presented three transit candidates for which we could not rule out the transiting planet model. Hence, at most we have detected 3 transiting hot Jupiters, although our analysis of these candidates shows that this is very unlikely. Follow-up observations will most likely show that our transit candidates are eclipsing binaries, which means that our transit survey will have produced a null result. From a simple signal-to-noise argument presented in Section 3.4.5, we expected to detect  $\sim 5$  transiting hot Jupiters based on a typical radius of  $\sim 1.4R_J$  and Solar neighbourhood frequency of  $\sim 1\%$  for Sun-like stars.

This analysis, however, has ignored/approximated many factors that may affect the accuracy of our estimate of the number of hot Jupiters that we expected to detect. These include:

1. Limb darkening effects which tend to make central eclipses deeper and grazing eclipses shallower.



2. The effect of orbital inclination on the transit signature of an extra-solar planet.
3. The distribution of our photometric data in time and the individual error bars on each measurement.
4. The effect of our modified transit detection algorithm on the S/N of a transit signal.
5. The minimum number of data points in-transit and out-of-transit required for a detection.
6. The rate of detections of spurious transit signals due to noise and/or systematic errors in the lightcurves (the false alarm rate).

The expected number of hot Jupiter detections and false alarms depend on the set  $Y$  of stars considered, the detection threshold  $S_{\min}^2$ , the extra-solar planet period  $P$  and the extra-solar planet radius  $R_p$ . We ignored these functional dependencies in our estimate of  $\sim 5$  expected hot Jupiter detections by considering a fixed detection threshold ( $S_{\min}^2 = 100$ ), period ( $P = 3.5$  d) and planetary radius ( $R_p = 1.4R_J$ ) for all stars in our sample. In order to improve our estimate of the number of hot Jupiters that we expected to detect and in order to estimate our false alarm rate, we have carried out Monte Carlo simulations on the lightcurves in our data set.

## 5.2 Detection Probabilities And False Alarm Rates

Consider an extra-solar planet of radius  $R_p$ , orbital period  $P$  and orbital inclination  $i$  with  $t_0$  as the time of central transit. Let the planet be in orbit around a star  $S$  of known mass and radius that has an associated lightcurve. Then we may calculate the predicted transit lightcurve of the planet (using the program *transitcurve.pro* - see Section 2.4.3) and add this signal into the observed lightcurve of the star. Consequently we may calculate the transit statistic  $S_{\text{tra}}^2$  (Section 4.1) using the values of  $t_0$  and  $\Delta t$  for each transit event, and

then evaluate the following detection function:

$$D_S(S_{\min}^2, N_{\min}, R_p, P, i, t_0) = \begin{cases} 1 & \text{if } S_{\text{tra}}^2 \geq S_{\min}^2 \text{ for at least } N_{\min} \text{ predicted transits} \\ 0 & \text{otherwise} \end{cases} \quad (5.1)$$

where  $S_{\min}^2$  is the transit statistic detection threshold. Using the same procedure as above, but without actually adding the predicted transit lightcurve into the observed lightcurve of the star, we may evaluate the false alarm function:

$$F_S(S_{\min}^2, N_{\min}, R_p, P, i, t_0) = \begin{cases} 1 & \text{if } S_{\text{tra}}^2 \geq S_{\min}^2 \text{ for at least } N_{\min} \text{ predicted transits} \\ 0 & \text{otherwise} \end{cases} \quad (5.2)$$

Since the transit statistic for the observed lightcurve with the injected transit is in general greater than or equal to the transit statistic for the observed lightcurve without the injected transit, it is clear that:

$$D_S(S_{\min}^2, N_{\min}, R_p, P, i, t_0) \geq F_S(S_{\min}^2, N_{\min}, R_p, P, i, t_0) \quad (5.3)$$

The parameters  $t_0$  and  $i$  in the detection function  $D_S$  are nuisance parameters since they do not reveal any useful physical information about the extra-solar planet itself. In order to remove these dimensions from the parameter space, we must multiply the detection function  $D_S$  by the corresponding joint probability distribution function (PDF)  $f(t_0, i)$ , and then integrate over the appropriate ranges. The parameter  $t_0$  lies in the range  $0 \leq t_0 \leq P$  and the parameter  $i$  lies in the range  $0^\circ \leq i \leq 90^\circ$ . Hence we have:

$$P(\text{det} | S, S_{\min}^2, N_{\min}, R_p, P) = \int_{i=0^\circ}^{i=90^\circ} \int_{t_0=0}^{t_0=P} f(t_0, i) D_S(S_{\min}^2, N_{\min}, R_p, P, i, t_0) dt_0 di \quad (5.4)$$

where  $P(\text{det} | S, S_{\min}^2, N_{\min}, R_p, P)$  is the detection probability for star  $S$  and  $f(t_0, i)$  is the joint PDF of  $t_0$  and  $i$ . We assume that the parameters  $t_0$  and  $i$  are independent and therefore:

$$f(t_0, i) = f(t_0)f(i) \quad (5.5)$$

We now assume that the time of central transit and the orbital inclination are random

(as in Section 2.4.1). This means that  $t_0$  is distributed uniformly between 0 and  $P$ :

$$\begin{aligned} P(0 \leq x \leq t_0) &= \frac{t_0}{P} \\ \therefore \frac{dP}{dt_0} &= \frac{1}{P} \\ \therefore f(t_0) &= \frac{1}{P} \quad \text{for } 0 \leq t_0 \leq P \end{aligned} \quad (5.6)$$

We may also use Equation [6] in Appendix A:

$$\begin{aligned} P(0 \leq x \leq i) &= 1 - \cos i \\ \therefore \frac{dP}{di} &= \sin i \\ \therefore f(i) &= \sin i \quad \text{for } 0^\circ \leq i \leq 90^\circ \end{aligned} \quad (5.7)$$

Substitute Equations 5.6 and 5.7 into Equation 5.5:

$$f(t_0, i) = \frac{\sin i}{P} \quad (5.8)$$

Now substitute Equation 5.8 into Equation 5.4:

$$P(\det | S, S_{\min}^2, N_{\min}, R_p, P) = \int_{i=0^\circ}^{i=90^\circ} \int_{t_0=0}^{t_0=P} \left( \frac{\sin i}{P} \right) D_S(S_{\min}^2, N_{\min}, R_p, P, i, t_0) dt_0 di \quad (5.9)$$

Using a parallel argument, we obtain an expression for the false alarm probability

$P(\text{fal} | S, S_{\min}^2, N_{\min}, R_p, P)$  as:

$$P(\text{fal} | S, S_{\min}^2, N_{\min}, R_p, P) = \int_{i=0^\circ}^{i=90^\circ} \int_{t_0=0}^{t_0=P} \left( \frac{\sin i}{P} \right) F_S(S_{\min}^2, N_{\min}, R_p, P, i, t_0) dt_0 di \quad (5.10)$$

### 5.3 Monte Carlo Simulations

We decided to take the Monte Carlo approach to evaluating the detection probabilities and false alarm rates, rather than attempting to numerically integrate Equations 5.9 and 5.10. In general, a Monte Carlo simulation attempts to estimate the required probability of an



event by selecting a large random sample from the parameter space as governed by the underlying PDF, and then calculating the fraction of the sample that satisfy the event criteria. The larger the sample, the more accurate the calculated probability. However, the size of the sample that may be selected and analysed is usually limited by available computing resources.

For each star  $S$ , we used the Monte Carlo method to calculate  $P(\text{det} | S, S_{\min}^2, N_{\min}, R_p, P)$  and  $P(\text{fal} | S, S_{\min}^2, N_{\min}, R_p, P)$  for a grid in  $S_{\min}^2, N_{\min}, R_p$  and  $P$  using the lightcurve data from the 2000-09 run. We chose to use a geometric sequence in  $S_{\min}^2$  from  $S_{\min}^2 = 5.6$  to  $S_{\min}^2 = 699$  with geometric factor 1.05. We chose  $N_{\min} \in \{1, 2, 3\}$  and  $R_p = 1.40R_J$ .

The grid for  $P$  should be fine enough that the difference in period  $\Delta P = P_{i+1} - P_i$  between two consecutive grid points  $P_i$  and  $P_{i+1}$  (where  $P_{i+1} > P_i$ ) is such that the difference in the number of cycles spanning the duration of the lightcurve is less than or equal to a fraction  $f_t$  of the transit duration (in cycle units). Let us denote the duration of the lightcurve by  $T$  and the transit duration by  $\Delta t$ . Expressing this condition in a mathematical form yields:

$$\frac{T}{P_i} - \frac{T}{P_{i+1}} \leq \frac{f_t \Delta t}{P_{i+1}} \quad (5.11)$$

Rearranging Equation 5.11 we get:

$$\frac{P_{i+1}}{P_i} \leq 1 + \frac{f_t \Delta t}{T} \quad (5.12)$$

From Equation 5.12 it is clear that the grid in  $P$  should be a geometric sequence with geometric factor less than or equal to  $1 + (f_t \Delta t / T)$ . We adopt  $f_t = 0.5$  in this analysis. For the 2000-09 run we have  $T = 10.4$  d and we take  $\Delta t \approx 2$  h for a typical transit duration which yields  $1 + (f_t \Delta t / T) \approx 1.004$ . Since we are interested in hot Jupiters, we used a grid in  $P$  as a geometric sequence from  $P = 1$  d to  $P = 10$  d with the geometric factor 1.004, which leads to 576 period grid points.

In our Monte Carlo simulations we carried out the following steps:

1. For each star  $S$ , planet radius  $R_p$  and orbital period  $P$  we carried out steps 2-5.
2. We selected a set  $X$  of  $N_{\text{MC}} = 1000$  planets of radius  $R_p$  with  $t_0$  and  $i$  drawn randomly for each planet from the PDFs in Equations 5.6 and 5.7 respectively. The value of the

orbital period  $P_{\text{curr}}$  for the current planet was drawn from a uniform distribution on the interval  $P/\sqrt{1.004} \leq P_{\text{curr}} \leq P\sqrt{1.004}$  in order to account for the planets which have a transit duration shorter than 2 h.

3. We calculated  $D_S(S_{\text{min}}^2, N_{\text{min}}, R_p, P, i, t_0)$  for each planet,  $S_{\text{min}}^2$  and  $N_{\text{min}}$ . We then obtained an estimate for  $P(\text{det} | S, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  from:

$$P(\text{det} | S, S_{\text{min}}^2, N_{\text{min}}, R_p, P) \approx \frac{1}{N_{\text{MC}}} \sum_X D_S(S_{\text{min}}^2, N_{\text{min}}, R_p, P, i, t_0) \quad (5.13)$$

4. We also calculated  $F_S(S_{\text{min}}^2, N_{\text{min}}, R_p, P, i, t_0)$  for each planet,  $S_{\text{min}}^2$  and  $N_{\text{min}}$ . We then obtained an estimate for  $P(\text{fal} | S, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  from:

$$P(\text{fal} | S, S_{\text{min}}^2, N_{\text{min}}, R_p, P) \approx \frac{1}{N_{\text{MC}}} \sum_X F_S(S_{\text{min}}^2, N_{\text{min}}, R_p, P, i, t_0) \quad (5.14)$$

5. Let us denote the uncertainty in  $P(\text{det} | S, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  and  $P(\text{fal} | S, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  as  $\sigma_{P_{\text{det}}}(S, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  and  $\sigma_{P_{\text{fal}}}(S, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  respectively. Then, assuming Poisson statistics, we have:

$$\sigma_{P_{\text{det}}}(S, S_{\text{min}}^2, N_{\text{min}}, R_p, P) = \frac{1}{N_{\text{MC}}} \sqrt{\sum_X D_S(S_{\text{min}}^2, N_{\text{min}}, R_p, P, i, t_0)} \quad (5.15)$$

$$\sigma_{P_{\text{fal}}}(S, S_{\text{min}}^2, N_{\text{min}}, R_p, P) = \frac{1}{N_{\text{MC}}} \sqrt{\sum_X F_S(S_{\text{min}}^2, N_{\text{min}}, R_p, P, i, t_0)} \quad (5.16)$$

In Section 5.1, we noted the factors which we ignored/approximated in our previous estimate of the number of hot Jupiters that we expected to detect. The Monte Carlo simulations take all of these factors into account when calculating the detection and false alarm probabilities. By explicitly using the function  $D_S(S_{\text{min}}^2, N_{\text{min}}, R_p, P, i, t_0)$ , we have immediately included the effects of points 1-5 in Section 5.1 and by using the corresponding false alarm function  $F_S(S_{\text{min}}^2, N_{\text{min}}, R_p, P, i, t_0)$ , we have taken care of point 6.

In Figure 5.1, we plot the detection probability and false alarm probability as functions of the transit statistic detection threshold (top) and period (bottom) for star 61377 during the 2000-09 run. This  $r' \approx 18.20$  mag G star has a mass, radius and distance of  $0.96M_{\odot}$ ,  $0.96R_{\odot}$  and 3152pc respectively. The detection probability decreases strongly as the detection threshold increases, as does the false alarm probability (Figure 5.1 top). For this

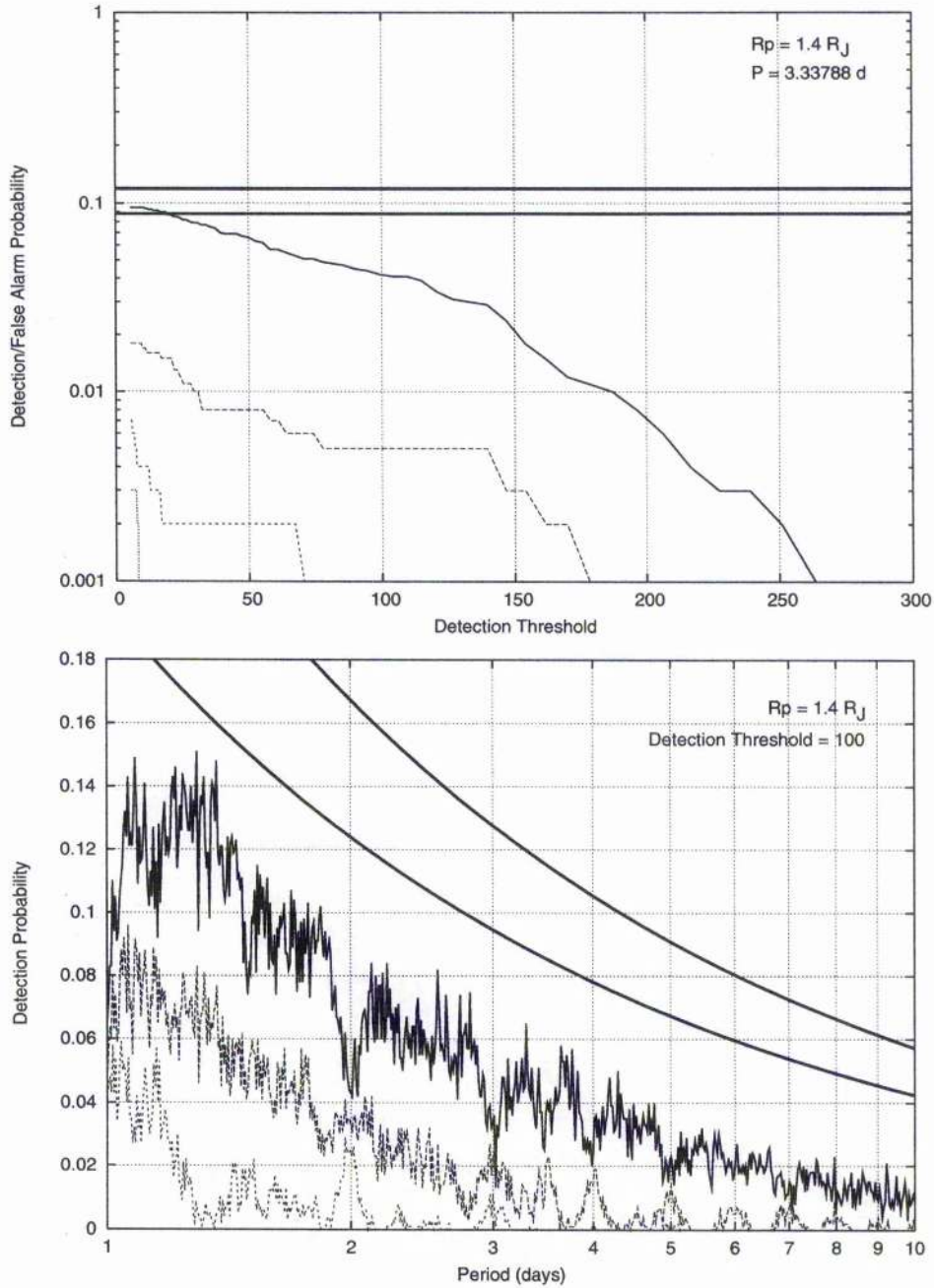


Figure 5.1: Detection probability as a function of  $S_{\min}^2$  for star 61377 during the 2000-09 run with  $P = 3.338$  d and  $R_p = 1.40R_J$  (**Top**) and as a function of  $P$  with  $S_{\min}^2 = 100$  and  $R_p = 1.40R_J$  (**Bottom**). The continuous line corresponds to  $N_{\min} = 1$ , the dashed line corresponds to  $N_{\min} = 2$  and the shorter dashed line corresponds to  $N_{\min} = 3$  in both diagrams. The false alarm probability as a function of  $S_{\min}^2$  for the same star and planet with  $N_{\min} = 1$  is shown by the dotted line in the top diagram. The false alarm probability is approximately zero for all  $P$  and  $N_{\min}$  in the bottom diagram. Star 61377 has 612 data points over 11 nights in its 2000-09 lightcurve with an RMS of  $\sim 0.010$  mag. It has a mass, radius and distance of  $0.96M_{\odot}$ ,  $0.96R_{\odot}$  and 3152pc respectively. In the top diagram, the upper and lower horizontal lines correspond to  $P_{\text{tra}}$  as calculated from Equation 2.9 and  $P_{\text{ann}}$  as calculated from Equation 2.10 respectively. In the bottom diagram, the upper and lower smooth continuous curves correspond to  $P_{\text{tra}}$  and  $P_{\text{ann}}$  respectively.

particular star, it can be seen that false alarms are very unlikely even at very low detection thresholds. The upper and lower horizontal continuous lines correspond to  $P_{\text{tra}} \approx 0.119$  as calculated from Equation 2.9 and  $P_{\text{ann}} \approx 0.088$  as calculated from Equation 2.10 respectively. These probabilities are independent of the detection threshold, and serve to highlight how the fraction of transiting planets that we are able to detect varies with detection threshold. For example, for  $S_{\text{min}}^2 = 100$  and  $N_{\text{min}} = 1$ , we recover  $\sim 30\%$  of the injected planets that transit star 61377.

In Figure 5.1 (bottom) one can see that the detection probability is highly dependent on the orbital period. For  $N_{\text{min}} = 1$ , orbital periods close to integer values tend to have lower detection probabilities since such periods are resonant with the observational gaps during the daytime. Conversely, orbital periods close to fractional values tend to have higher detection probabilities since such periods cover a greater range of orbital phases. For example, periods close to  $\sim 3.0$  d have a detection probability of  $\sim 0.02$  whereas periods close to  $\sim 3.3$  d have a detection probability of  $\sim 0.06$  for this particular star. Figure 5.1 also shows that as you increase the number of recovered transits required for a detection, the detection probability decreases rapidly. The upper and lower continuous curves correspond to  $P_{\text{tra}}$  as calculated from Equation 2.9 and  $P_{\text{ann}}$  as calculated from Equation 2.10 respectively. From these equations, and Kepler's third law (Equation 1.8), it is clear that  $P_{\text{tra}} \propto P^{-2/3}$  and  $P_{\text{ann}} \propto P^{-2/3}$ .

The calculations of the detection probabilities and false alarm rates were carried out using the CONDOR<sup>1</sup> workload management system. Each lightcurve requires  $\sim 30$  minutes of computing time on a 1.0Ghz machine. That is a total of  $\sim 1.9$  CPU-years required to analyse our  $\sim 33000$  lightcurves. One could nearly write a thesis in that time! CONDOR distributes the workload in batches to computers linked to the system, and when an individual computer is idle, it executes the current batch. We sent each lightcurve as a batch job to the CONDOR pools at St Andrews (47 CPUs) and the Instituto de Astrofísica de Canarias (91 CPUs). The calculations were finished in just under three weeks since we were limited to 30 IDL licenses at St Andrews and 20 IDL licenses at the IAC.

<sup>1</sup><http://www.cs.wisc.edu/condor>

Table 5.1: The different subsets of stars used when calculating the number of expected transiting planets and false alarms.

$Y$	No. Of Stars	Mass Range
All Stars	32027	$0.08M_{\odot} \leq M_* \leq 1.40M_{\odot}$
Late F Stars	3129	$1.05M_{\odot} < M_* \leq 1.40M_{\odot}$
G Stars	7423	$0.80M_{\odot} < M_* \leq 1.05M_{\odot}$
K Stars	15381	$0.50M_{\odot} < M_* \leq 0.80M_{\odot}$
M Stars	6094	$0.08M_{\odot} \leq M_* \leq 0.50M_{\odot}$

#### 5.4 Number Of Expected Transiting Planets

Assuming that each star has one planet of radius  $R_p$  and period  $P$ , then the number of expected transiting planets  $N_{\text{det}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  as a function of star type  $Y$ ,  $S_{\text{min}}^2$ ,  $N_{\text{min}}$ ,  $R_p$  and  $P$  is simply the sum of the detection probabilities for all stars of the required type:

$$N_{\text{det}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P) = \sum_{S \in Y} P(\text{det} | S, S_{\text{min}}^2, N_{\text{min}}, R_p, P) \quad (5.17)$$

Similarly, the number of expected false alarms  $N_{\text{fal}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  is given by:

$$N_{\text{fal}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P) = \sum_{S \in Y} P(\text{fal} | S, S_{\text{min}}^2, N_{\text{min}}, R_p, P) \quad (5.18)$$

The uncertainty in  $N_{\text{det}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  and  $N_{\text{fal}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$ , denoted by  $\sigma_{N_{\text{det}}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  and  $\sigma_{N_{\text{fal}}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  respectively, may be derived from Equations 5.13 through 5.18 as:

$$\sigma_{N_{\text{det}}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P) = \sqrt{\frac{1}{N_{\text{MC}}} N_{\text{det}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P)} \quad (5.19)$$

$$\sigma_{N_{\text{fal}}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P) = \sqrt{\frac{1}{N_{\text{MC}}} N_{\text{fal}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P)} \quad (5.20)$$

We are interested in the number of expected transiting planets (and false alarms) for stars of different masses or, equivalently, spectral types. To facilitate this analysis we consider 5 different sets of stars. The first set is the set of all stars with a lightcurve from



the 2000-09 run, and a derived mass, radius and distance. This set includes a total of 32027 stars. The other sets are mutually exclusive subsets of this set consisting of late F stars, G stars, K stars and M stars respectively. Table 5.1 shows the number of stars in each set and the mass/spectral type ranges to which they correspond. The mass ranges for the various spectral types are taken from Lang (1992).

In Figure 5.2, we plot the number of expected transiting planets and false alarms for all stars as functions of the transit statistic detection threshold for  $P = 3.338$  d (top) and for  $P = 3.008$  d (bottom). In both plots the continuous curves represent  $N_{\text{det}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  for  $N_{\text{min}} = 1$  (upper curve),  $N_{\text{min}} = 2$  (middle curve) and  $N_{\text{min}} = 3$  (lower curve). Similarly, the dashed curves represent  $N_{\text{fal}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P)$  for  $N_{\text{min}} = 1$  (upper curve),  $N_{\text{min}} = 2$  (middle curve) and  $N_{\text{min}} = 3$  (lower curve). It is interesting to note how the small change in period from  $P \approx 3.0$  d to  $P \approx 3.3$  d can seriously affect the number of expected transiting planets (and false alarms) in different ways for different values of  $N_{\text{min}}$ . For instance, when  $N_{\text{min}} = 1$ , we have:

$$N_{\text{det}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, 3.3 \text{ d}) > N_{\text{det}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, 3.0 \text{ d})$$

and:

$$N_{\text{fal}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, 3.3 \text{ d}) \approx N_{\text{fal}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, 3.0 \text{ d})$$

However, when  $N_{\text{min}} = 2$ , we have:

$$N_{\text{det}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, 3.3 \text{ d}) < N_{\text{det}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, 3.0 \text{ d})$$

and:

$$N_{\text{fal}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, 3.3 \text{ d}) < N_{\text{fal}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, 3.0 \text{ d})$$

This is easily explained by the fact that the phase coverage of the observations with  $N_{\text{min}} = 1$  is better for  $P = 3.3$  d than for  $P = 3.0$  d and conversely, the phase coverage of the observations with  $N_{\text{min}} = 2$  is better for  $P = 3.0$  d than for  $P = 3.3$  d.

One may also see from Figure 5.2 that  $N_{\text{fal}}(Y, S_{\text{min}}^2, N_{\text{min}}, R_p, P) \approx 0.3 < 1$  for both periods at our chosen transit statistic detection threshold  $S_{\text{min}}^2 = 100$  and  $N_{\text{min}} = 1$  (see Section 4.1). This is important because it means that our survey is unlikely to yield any



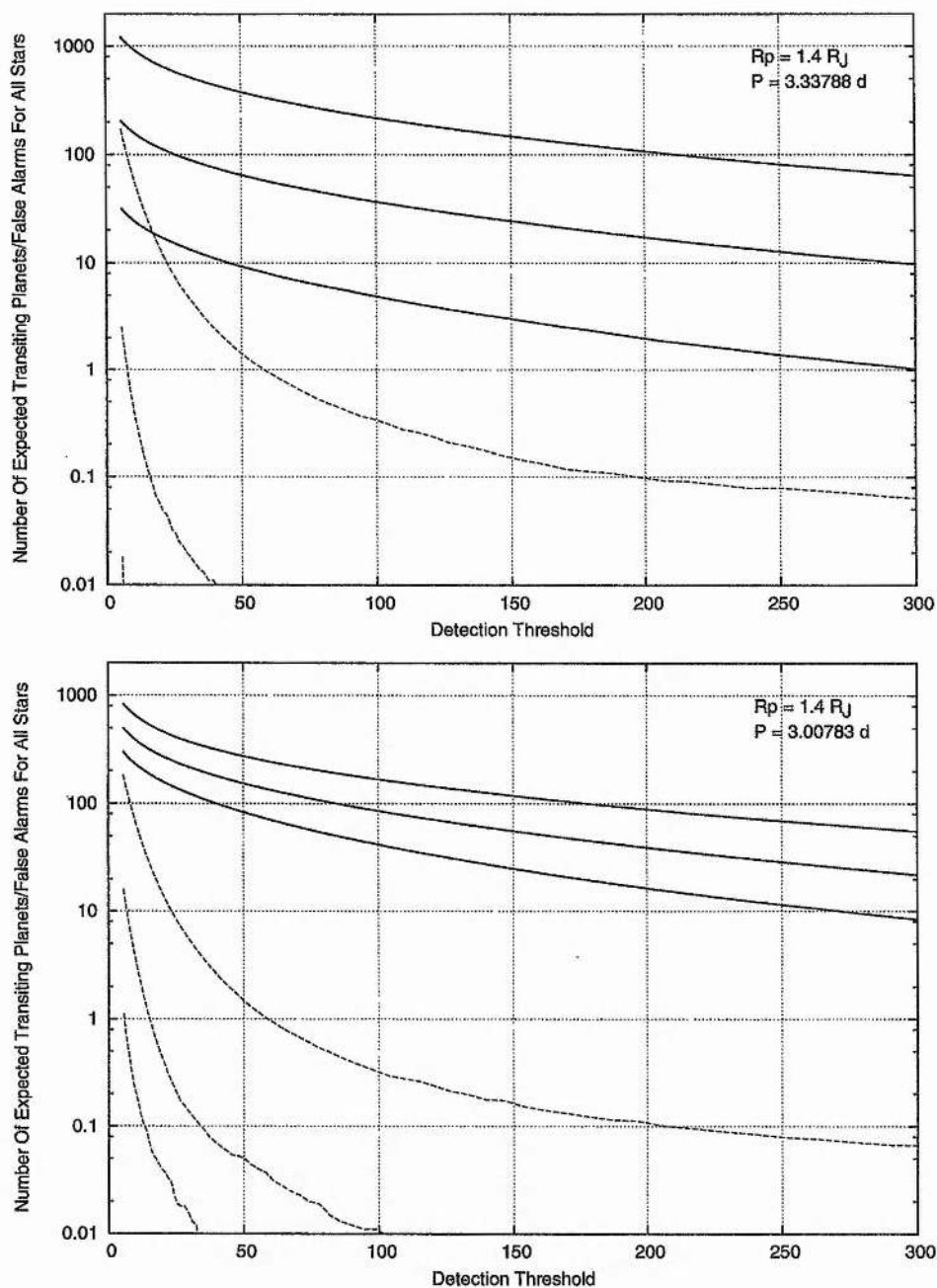


Figure 5.2: Number of expected transiting planets (continuous curves) and false alarms (dashed curves) for all stars as functions of  $S_{\min}^2$  with  $R_p = 1.40 R_J$ . The **top** diagram corresponds to  $P = 3.338 \text{ d}$  and the **bottom** diagram corresponds to  $P = 3.008 \text{ d}$ . In both diagrams the continuous curves represent  $N_{\text{det}}(Y, S_{\min}^2, N_{\min}, R_p, P)$  for  $N_{\min} = 1$  (upper curve),  $N_{\min} = 2$  (middle curve) and  $N_{\min} = 3$  (lower curve). Similarly, the dashed curves represent  $N_{\text{fal}}(Y, S_{\min}^2, N_{\min}, R_p, P)$  for  $N_{\min} = 1$  (upper curve),  $N_{\min} = 2$  (middle curve) and  $N_{\min} = 3$  (lower curve).

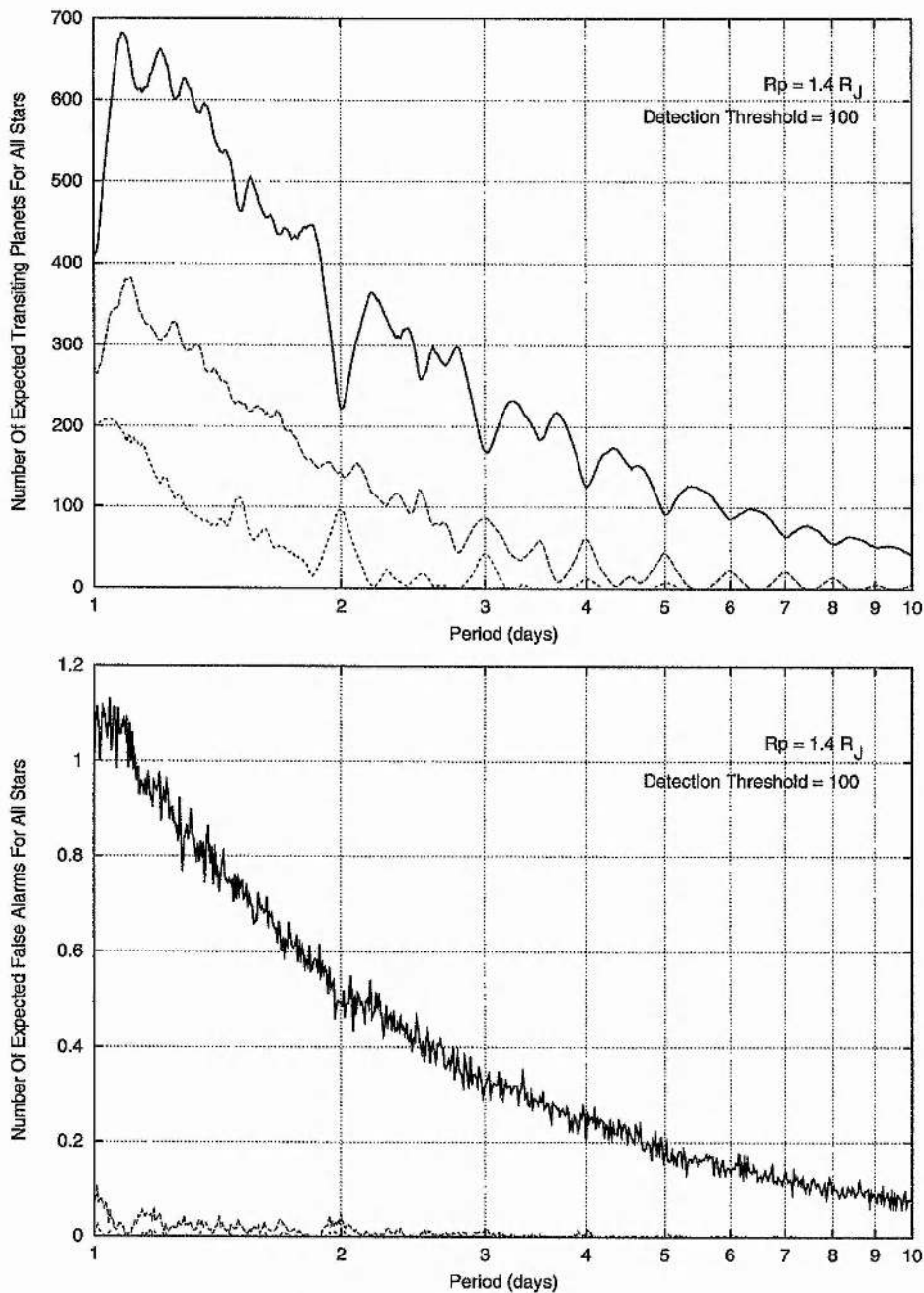


Figure 5.3: Number of expected transiting planets (Top) and the number of expected false alarms (Bottom) for all stars as functions of  $P$  for  $R_p = 1.40R_J$  and  $S_{\min}^2 = 100$ . In both diagrams, the continuous curve corresponds to  $N_{\min} = 1$ , the dashed curve to  $N_{\min} = 2$  and the shorter dashed curve to  $N_{\min} = 3$ .

transit candidates that we conclude are transiting planets when in fact they are not and thus our choice of such a high detection threshold is justified.

In Figure 5.3, we plot the number of expected transiting planets (top) and the number of expected false alarms (bottom) for all stars as functions of the period with  $R_p = 1.40R_J$  and  $S_{\min}^2 = 100$ . Again note the strong dependence of these quantities on the period. The bottom diagram clearly shows that, by increasing  $N_{\min}$  from 1 to 2, the number of expected false alarms is effectively reduced to zero for all  $P$ . However, introducing this extra constraint for a transit detection more than halves the expected yield of planets from the survey (Figure 5.3 top). By choosing  $N_{\min} = 1$  in Section 4.1, we have opted to risk the possibility of false alarms in exchange for more transit detections.

### 5.5 Number Of Expected Transiting Hot Jupiters

As we have seen in Section 5.4, the number of expected transiting planets is highly dependent on the orbital period. A slight change in the period of the planet considered can change  $N_{\det}(Y, S_{\min}^2, N_{\min}, R_p, P)$  by up to a factor of 2 (Figure 5.3 top). Instead of considering a specific period, we may consider a range of periods that fit with the type of planet that we are interested in, making some assumption about the underlying PDF.

For hot Jupiters we have a PDF for the period  $f(P)$  such that  $f(P) \propto P^{-1}$  as revealed by the radial velocity surveys (see Section 1.2.3). We consider three sets of planets defined by their assigned period ranges. These sets are the very hot Jupiters with  $1d \leq P \leq 3d$ , the shorter period hot Jupiters with  $3d \leq P \leq 5d$  and the hot Jupiters with  $1d \leq P \leq 10d$ . The period ranges (and planet types) are arbitrary but consistent with the known short period extra-solar planets.

The period PDF  $f(P)$  for a set of planets with periods in the range  $P_1 \leq P \leq P_M$  (where  $M \in \mathbb{N}$ ) such that  $f(P) \propto P^{-1}$  is given by:

$$f(P) = \frac{1}{P(\ln P_M - \ln P_1)} \quad \text{for } P_1 \leq P \leq P_M \quad (5.21)$$

The detection probability  $P(\det | S, S_{\min}^2, N_{\min}, R_p, P_1, P_M)$  for the set of planets with the period PDF from Equation 5.21 is then found by integrating the product of  $f(P)$  and

$P(\det | S, S_{\min}^2, N_{\min}, R_p, P)$  over the period range  $P_1 \leq P \leq P_M$ :

$$P(\det | S, S_{\min}^2, N_{\min}, R_p, P_1, P_M) = \int_{P_1}^{P_M} \left( \frac{1}{P(\ln P_M - \ln P_1)} \right) P(\det | S, S_{\min}^2, N_{\min}, R_p, P) dP \quad (5.22)$$

Introducing the change of variable  $Z = \ln P$  into Equation 5.22 yields:

$$P(\det | S, S_{\min}^2, N_{\min}, R_p, P_1, P_M) = \left( \frac{1}{\ln P_M - \ln P_1} \right) \int_{\ln P_1}^{\ln P_M} P(\det | S, S_{\min}^2, N_{\min}, R_p, P) dZ \quad (5.23)$$

Making a discrete approximation to the integral in Equation 5.23 gives:

$$P(\det | S, S_{\min}^2, N_{\min}, R_p, P_1, P_M) \approx \left( \frac{1}{\ln P_M - \ln P_1} \right) \sum_{P \in \{P_1, P_2, \dots, P_M\}} P(\det | S, S_{\min}^2, N_{\min}, R_p, P) dZ_i \quad (5.24)$$

In our Monte Carlo simulations we chose to calculate  $P(\det | S, S_{\min}^2, N_{\min}, R_p, P)$  for a grid in  $P$  consisting of a geometric sequence from  $P_1 = 1.0$  d to  $P_M = 10.0$  d with geometric factor  $k = 1.004$  resulting in  $M = 576$  period grid points. Denoting a general period grid point by  $P = P_i$  for  $i \in \{1, 2, \dots, M\}$ , we have:

$$\begin{aligned} P_{i+1} &= kP_i \\ \therefore \ln P_{i+1} &= \ln P_i + \ln k \end{aligned} \quad (5.25)$$

Hence, for this period grid,  $dZ_i = \ln k$  for all  $i \in \{1, 2, \dots, M\}$ . Using this fact in Equation 5.24 yields:

$$P(\det | S, S_{\min}^2, N_{\min}, R_p, P_1, P_M) \approx \left( \frac{\ln k}{\ln P_M - \ln P_1} \right) \sum_{P \in \{P_1, P_2, \dots, P_M\}} P(\det | S, S_{\min}^2, N_{\min}, R_p, P) \quad (5.26)$$

The uncertainty in  $P(\det | S, S_{\min}^2, N_{\min}, R_p, P_1, P_M)$ , denoted by  $\sigma_{P_{\det}}(S, S_{\min}^2, N_{\min}, R_p, P_1, P_M)$ , may be derived from Equations 5.13, 5.15 and 5.26 as:

$$\sigma_{P_{\det}}(S, S_{\min}^2, N_{\min}, R_p, P_1, P_M) = \sqrt{\left( \frac{\ln k}{N_{MC}(\ln P_M - \ln P_1)} \right) P(\det | S, S_{\min}^2, N_{\min}, R_p, P_1, P_M)} \quad (5.27)$$

Assuming that each star has one planet of radius  $R_p$  and period  $P$  distributed with PDF  $f(P)$  from Equation 5.21 in the range  $P_1 \leq P \leq P_M$ , then the number of expected

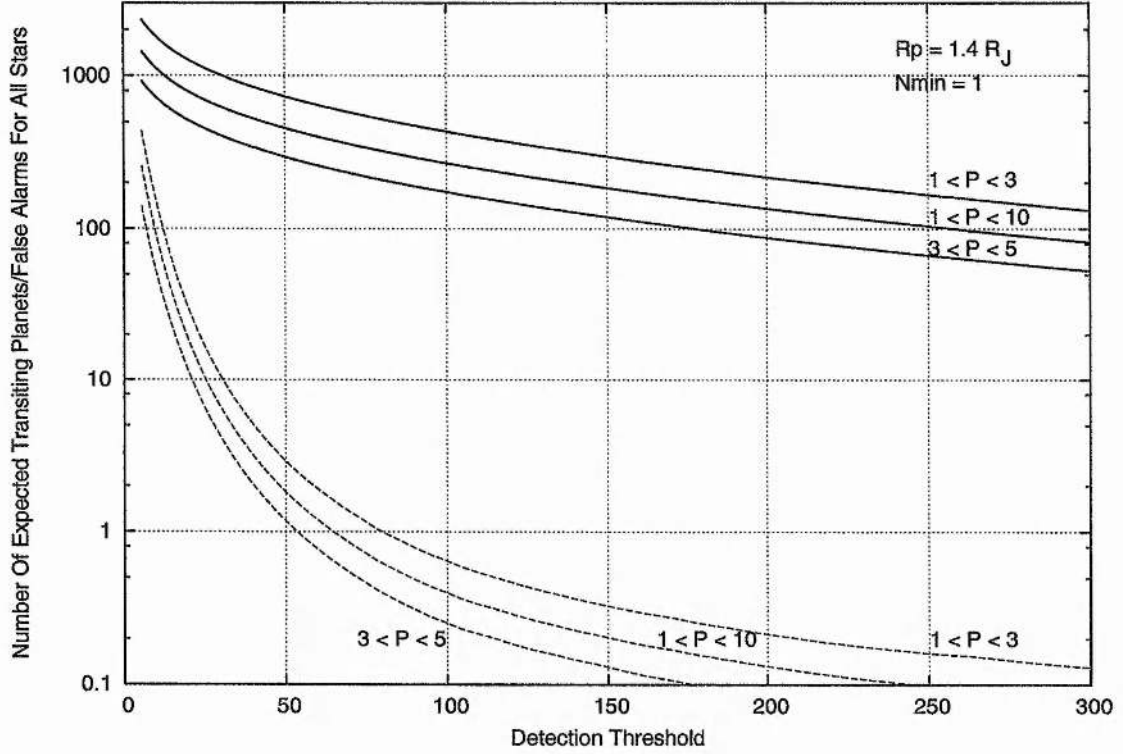


Figure 5.4: Number of expected transiting planets (continuous curves) and false alarms (dashed curves) for all stars as functions of  $S_{\min}^2$  with  $N_{\min} = 1$  and  $R_p = 1.40R_J$ . The period ranges that correspond to each curve are marked on the diagram.

planets  $N_{\det}(Y, S_{\min}^2, N_{\min}, R_p, P_1, P_M)$  as a function of star type  $Y$ ,  $S_{\min}^2$ ,  $N_{\min}$ ,  $R_p$ ,  $P_1$  and  $P_M$  is simply the sum of the detection probabilities for all stars of the required type:

$$N_{\det}(Y, S_{\min}^2, N_{\min}, R_p, P_1, P_M) = \sum_{S \in Y} P(\det | S, S_{\min}^2, N_{\min}, R_p, P_1, P_M) \quad (5.28)$$

The uncertainty in  $N_{\det}(Y, S_{\min}^2, N_{\min}, R_p, P_1, P_M)$ , denoted by

$\sigma_{N_{\det}}(Y, S_{\min}^2, N_{\min}, R_p, P_1, P_M)$ , may be derived from Equations 5.27 and 5.28 as:

$$\sigma_{N_{\det}}(Y, S_{\min}^2, N_{\min}, R_p, P_1, P_M) = \sqrt{\left( \frac{\ln k}{N_{MC}(\ln P_M - \ln P_1)} \right) N_{\det}(Y, S_{\min}^2, N_{\min}, R_p, P_1, P_M)} \quad (5.29)$$

Using a similar set of arguments to those presented above, we may derive the corresponding set of equations relating to the false alarm probability, the number of expected

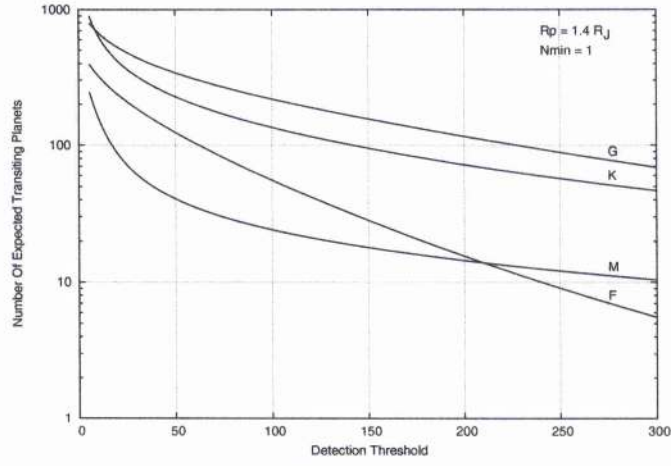
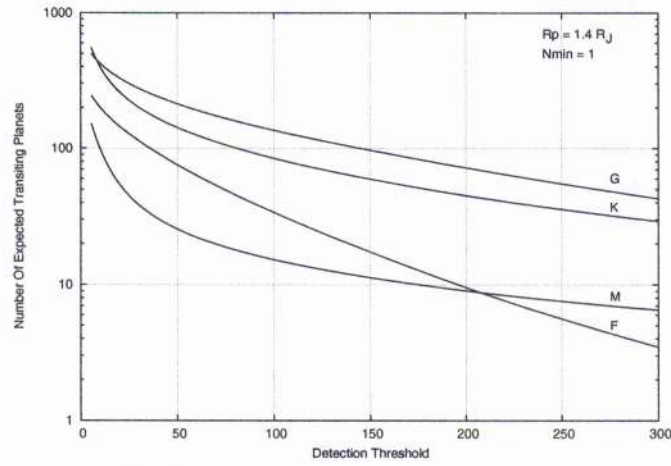
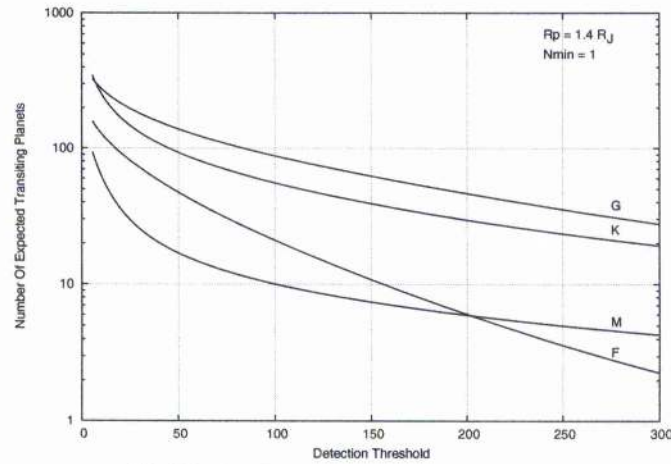
(a) The very hot Jupiters with  $1d \leq P \leq 3d$ .(b) The hot Jupiters with  $1d \leq P \leq 10d$ .(c) The hot Jupiters with  $3d \leq P \leq 5d$ .

Figure 5.5: Number of expected transiting planets as a function of  $S_{\min}^2$  with  $N_{\min} = 1$  and  $R_p = 1.40 R_J$ . The star types that correspond to each curve are marked on each diagram.



false alarms and the associated uncertainties:

$$P(\text{fal} | S, S_{\min}^2, N_{\min}, R_p, P_1, P_M) \approx \left( \frac{\ln k}{\ln P_M - \ln P_1} \right) \sum_{P \in \{P_1, P_2, \dots, P_M\}} P(\text{fal} | S, S_{\min}^2, N_{\min}, R_p, P) \quad (5.30)$$

$$\sigma_{P_{\text{fal}}}(S, S_{\min}^2, N_{\min}, R_p, P_1, P_M) = \sqrt{\left( \frac{\ln k}{N_{\text{MC}}(\ln P_M - \ln P_1)} \right) P(\text{fal} | S, S_{\min}^2, N_{\min}, R_p, P_1, P_M)} \quad (5.31)$$

$$N_{\text{fal}}(Y, S_{\min}^2, N_{\min}, R_p, P_1, P_M) = \sum_{S \in Y} P(\text{fal} | S, S_{\min}^2, N_{\min}, R_p, P_1, P_M) \quad (5.32)$$

$$\sigma_{N_{\text{fal}}}(Y, S_{\min}^2, N_{\min}, R_p, P_1, P_M) = \sqrt{\left( \frac{\ln k}{N_{\text{MC}}(\ln P_M - \ln P_1)} \right) N_{\text{fal}}(Y, S_{\min}^2, N_{\min}, R_p, P_1, P_M)} \quad (5.33)$$

Table 5.2: Results of the Monte Carlo simulations for  $R_p = 1.40R_J$  and  $S_{\min}^2 = 100$ . The number in brackets indicates the uncertainty on the last decimal place. The last two columns indicate the derived upper limits on  $f_p$  at the significance level  $\alpha$  for  $N_{\min} = 1$ .

Y	No. Of Stars	P (days)	$N_{\min}^{\text{det}} = 1$ with $N_{\min}^{\text{det}} = 1$	$N_{\min}^{\text{fal}} = 1$ with $N_{\min}^{\text{fal}} = 1$	$N_{\min}^{\text{det}} = 2$ with $N_{\min}^{\text{det}} = 2$	$N_{\min}^{\text{fal}} = 2$ with $N_{\min}^{\text{fal}} = 2$	$f_p$ at $\alpha = 0.01$	$f_p$ at $\alpha = 0.05$
All Stars	32027	3.00783	168.7(4)	0.325(18)	86.4(3)	0.011(3)	2.7%	1.8%
All Stars	32027	3.33788	219.4(5)	0.343(19)	36.9(2)	0.000	2.1%	1.4%
All Stars	32027	$1d < P < 3d$	436.49(4)	0.656(2)	199.27(3)	0.018(1)	1.0%	0.69%
All Stars	32027	$3d < P < 5d$	175.30(4)	0.257(1)	34.48(2)	0.003(1)	2.6%	1.7%
All Stars	32027	$1d < P < 10d$	271.24(2)	0.407(1)	105.13(1)	0.009(1)	1.7%	1.1%
Late F Stars	3129	3.00783	23.3(2)	0.092(10)	8.3(1)	0.005(2)	20%	13%
Late F Stars	3129	3.33788	25.3(2)	0.093(10)	3.8(1)	0.000	18%	12%
Late F Stars	3129	$1d < P < 3d$	56.17(1)	0.187(1)	17.58(1)	0.004(1)	8.2%	5.3%
Late F Stars	3129	$3d < P < 5d$	21.18(1)	0.073(1)	2.89(1)	0.001(1)	22%	14%
Late F Stars	3129	$1d < P < 10d$	34.21(1)	0.113(1)	9.22(1)	0.002(1)	13%	8.7%
G Stars	7423	3.00783	83.2(3)	0.133(12)	45.6(2)	0.004(2)	5.5%	3.6%
G Stars	7423	3.33788	111.4(3)	0.144(12)	19.2(1)	0.000	4.1%	2.7%
G Stars	7423	$1d < P < 3d$	219.82(3)	0.261(1)	106.85(2)	0.007(1)	2.1%	1.4%
G Stars	7423	$3d < P < 5d$	88.27(3)	0.110(1)	18.52(1)	0.001(1)	5.2%	3.4%
G Stars	7423	$1d < P < 10d$	136.63(2)	0.166(1)	56.37(1)	0.004(1)	3.4%	2.2%
K Stars	15381	3.00783	52.8(2)	0.094(10)	27.4(2)	0.002(1)	8.7%	6.7%
K Stars	15381	3.33788	69.9(3)	0.104(10)	11.9(1)	0.000	6.6%	4.3%
K Stars	15381	$1d < P < 3d$	136.13(2)	0.200(1)	63.11(2)	0.007(1)	3.4%	2.2%
K Stars	15381	$3d < P < 5d$	55.76(2)	0.073(1)	11.07(1)	0.001(1)	8.2%	5.4%
K Stars	15381	$1d < P < 10d$	85.12(1)	0.122(1)	33.35(1)	0.004(1)	5.4%	3.5%
M Stars	6094	3.00783	9.5(1)	0.001(1)	5.0(1)	0.000	48%	31%
M Stars	6094	3.33788	12.9(1)	0.002(1)	2.0(1)	0.000	36%	23%
M Stars	6094	$1d < P < 3d$	24.37(1)	0.008(1)	11.72(1)	0.000	19%	12%
M Stars	6094	$3d < P < 5d$	10.08(1)	0.002(1)	2.00(1)	0.000	46%	30%
M Stars	6094	$1d < P < 10d$	15.29(1)	0.005(1)	6.18(1)	0.000	30%	19%

In Table 5.2 we present the results of the Monte Carlo simulations for various values/ranges of the relevant parameters. The table details the number of expected transiting planets and false alarms for various pertinent sets of stars, period ranges, and values of  $N_{\min}$  as calculated for  $R_p = 1.40R_J$  and  $S_{\min}^2 = 100$ .

In Figure 5.4 we plot the number of expected transiting planets (continuous curves) and false alarms (dashed curves) for all stars as functions of the transit statistic detection threshold. The period ranges that correspond to each curve are marked on the diagram. One can see that we expected to detect  $\sim 436$  1d to 3d very hot Jupiters,  $\sim 175$  3d to 5d hot Jupiters and  $\sim 271$  1d to 10d hot Jupiters at a detection threshold of  $S_{\min}^2 = 100$  with  $N_{\min} = 1$  (Table 5.2) based on the assumption that each star has a single such planet of the specified type.

However, in reality, only a fraction  $f_p$  of the stars considered will harbour a planet of the specified type, and hence we must correct our calculations of the number of expected transiting planets by this factor. We refer to  $f_p$  as the planet fraction. Since  $f_p$  is an unknown quantity that we would like to estimate, we may use the fact that our transit survey has most likely produced a null result (although this is still to be confirmed) and place a significant upper limit on  $f_p$ . First of all we make the assumption that the number  $X$  of transiting planet detections has a Poisson distribution with expected value  $E(X)$  given by:

$$E(X) = f_p N_{\det} \quad (5.34)$$

The Poisson distribution is defined by:

$$P(X = x) = \frac{(E(X))^x}{x!} e^{-E(X)} \quad \text{for } x \in \mathbb{N}_0 \quad (5.35)$$

For a null result,  $x = 0$ . Using this fact and substituting Equation 5.34 into Equation 5.35 we get:

$$P(X = 0) = e^{-f_p N_{\det}} \quad (5.36)$$

In order to obtain an upper limit on  $f_p$  at the significance level  $\alpha$  we require:

$$P(X = 0) \leq \alpha \quad (5.37)$$

Using Equation 5.36 in Equation 5.37 we get:

$$e^{-f_p N_{\text{det}}} \leq \alpha \quad (5.38)$$

$$\therefore f_p \leq -\frac{\ln \alpha}{N_{\text{det}}} \quad (5.39)$$

The values of  $f_p$  that we derive in this manner for  $\alpha = 0.01$  and  $\alpha = 0.05$  are shown in the final two columns of Table 5.2.

The results from Table 5.2 and Figure 5.4 indicate that, for all the stars in our sample and at a significance level of 1%, we may place an upper limit of 1.0% on the 1d to 3d very hot Jupiter fraction, an upper limit of 2.6% on the 3d to 5d hot Jupiter fraction and an upper limit of 1.7% on the 1d to 10d hot Jupiter fraction, based on the assumption that such planets have a typical radius of  $\sim 1.40R_J$ .

Figure 5.5 shows the number of expected transiting planets as a function of the transit statistic detection threshold for the different period ranges and stellar types. Figure 5.5(a) corresponds to the period range  $1d \leq P \leq 3d$ , Figure 5.5(b) corresponds to the period range  $1d \leq P \leq 10d$  and Figure 5.5(c) corresponds to the period range  $3d \leq P \leq 5d$ . Each curve corresponds to a different star type as defined in Table 5.1. For our survey with  $S_{\text{min}}^2 = 100$  and  $N_{\text{min}} = 1$ , the best limits that we are able to place on the planet fraction are for the G stars in our sample. For these stars, at a significance level of 1%, we constrain  $f_p \leq 2.1\%$  for the 1d to 3d very hot Jupiters,  $f_p \leq 5.2\%$  for the 3d to 5d hot Jupiters and  $f_p \leq 3.4\%$  for the 1d to 10d hot Jupiters.

## 5.6 Discussion

We may now attempt to answer the question posed in Section 1.2.4, "What fraction of Sun-like stars have planets, and how does it depend on the host star properties?". Instead of providing an estimate of the planet fraction as a function of planet type and star type, we have managed to derive relatively stringent upper limits on the abundance of planets for the field of NGC 7789. Figure 5.6 shows a plot of the upper limit on the planet fraction  $f_p$  at a significance level of 1% as a function of star type for the 1d to 3d very hot Jupiters

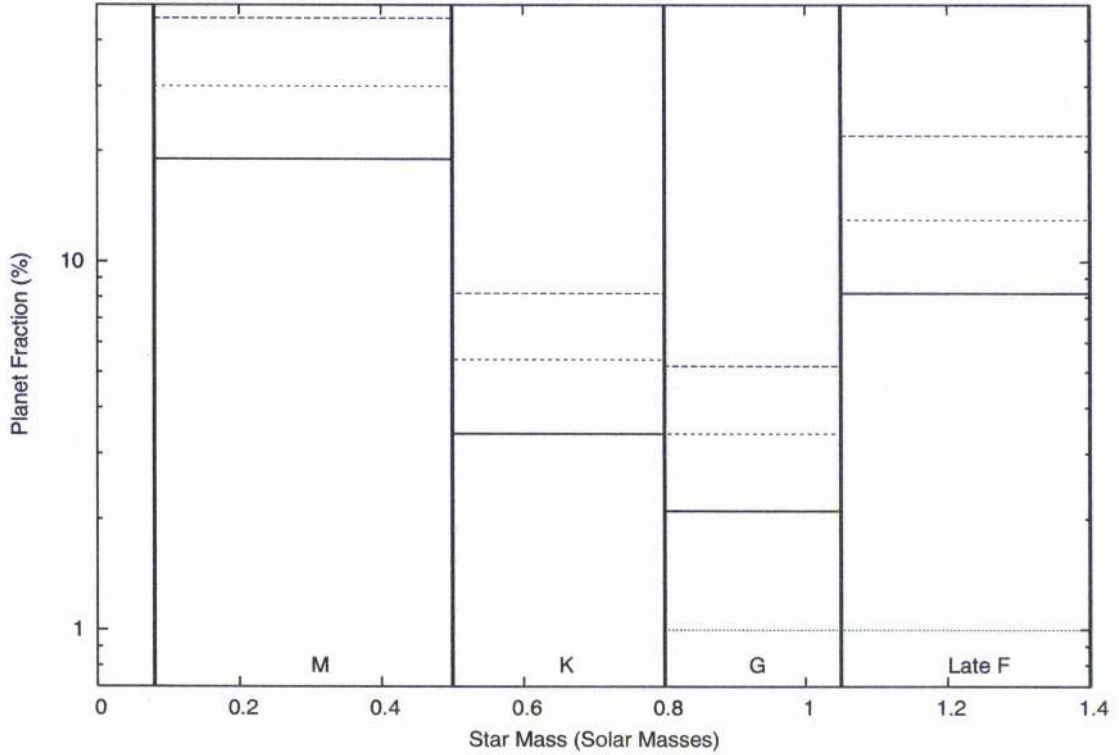


Figure 5.6: The upper limit on  $f_p$  against star type for the 1d to 3d very hot Jupiters (continuous line), the 3d to 5d hot Jupiters (dashed line) and the 1d to 10d hot Jupiters (shorter dashed line). The dotted line shows the estimate of the 3d to 5d hot Jupiter fraction for Solar neighbourhood stars from Butler et al. (2000).

(continuous line), the 3d to 5d hot Jupiters (dashed line) and the 1d to 10d hot Jupiters (shorter dashed line). We also show the estimate by Butler et al. (2000) that  $\sim 1\%$  of nearby Sun-like stars (late F and G dwarfs) host a 3d to 5d hot Jupiter (dotted line), as derived from radial velocity observations.

It is interesting to note that although the K stars are the most numerous in our star sample (15381 stars), it is the 7423 G stars that produce the largest number of expected transiting planets, and therefore the strictest limits on  $f_p$ . This is due to the fact that the G stars are generally brighter than the K stars in our sample, and therefore the corresponding gain in accuracy of the photometric measurements outweighs the smaller number of stars for which we can search for transits and the smaller transit signal for a given planetary

radius.

We may compare our results directly with those of Butler et al. (2000) by considering the late F and G stars in our sample and the corresponding number of expected transiting planets for the 3d to 5d hot Jupiters. We expected to detect  $\sim 21$  such planets around the late F stars in our sample, and  $\sim 88$  such planets around the G stars in our sample, a total of  $\sim 109$  expected 3d to 5d hot Jupiters. This places an upper limit on  $f_p$  of  $\sim 4.2\%$  at the 1% significance level for these types of star and planet. This is consistent with the value derived by Butler et al. (2000) of  $f_p \approx 1\%$  and demonstrates with confidence that the hot Jupiter fraction for Sun-like stars in this field may not be more than a factor of  $\sim 4$  times greater than that for the Solar neighbourhood. In our Monte Carlo simulations we have only considered the lightcurve data from the 2000-09 run. By including the lightcurve data from the 1999-07 run in any future simulations, where we also looked for transits, we will be able to place better limits on the hot Jupiter fraction in this field.

Our previous estimate of  $\sim 5$  transiting hot Jupiters from Section 3.4.5, based on  $f_p = 1.0\%$ , is clearly an over estimate. The Monte Carlo simulations indicate that for  $f_p = 1.0\%$  we should have expected to detect  $\sim 2$  3d to 5d hot Jupiters from all the stars in our sample (Table 5.2). The previous estimate fails in that it does not take into account any of the factors mentioned in Section 5.1 that affect the sensitivity of a transit survey. We note that by using the actual lightcurves of the stars in our sample we account exactly for the distribution of our observations in time, but introduce some extra false alarms from the existence of variable star lightcurves in the sample. This means that we tend to slightly over estimate the false alarm rate which in turn leads us to choose a slightly higher detection threshold than is necessary.

We conclude that our transit survey from the 2000-09 run alone has just about reached the sensitivity required to detect a few hot Jupiters if the abundance of such planets is similar to that of the Solar neighbourhood.



## 5.7 Summary

This chapter has been used to present the ideas of detection probability and false alarm probability for photometric observations of a star with an extra-solar planet. For each star in our sample we have used Monte Carlo simulations to evaluate these probabilities for a grid in planet period, planet radius, detection threshold and the minimum number of recovered transits required for a detection. Consequently we have derived an accurate estimate for the number of expected transiting planets and false alarms from our transit survey as a function of star and planet type. This has allowed us to place corresponding limits at the significance level of 1% on the hot Jupiter fraction for the different types of stars in this field under the assumption that our survey has produced a null result.

# 6

---

## Conclusions

In the search for our transit candidates we have developed an accurate, efficient and fast photometry pipeline employing the technique of difference image analysis. Raw data from the telescope are processed by the pipeline with minimal user input in order to directly produce lightcurves and colour magnitude diagrams. This is especially important considering the high quantity of data that may arise from a transit survey.

Our analysis of the colour magnitude diagrams by including the treatment of extinction for the open cluster NGC 7789 has allowed us to assign a model-dependent mass, radius and distance to each star. Such information is vital in the subsequent analysis of the transit candidates since it allows a direct estimate of the companion radius. We detected 24 transit candidates which warranted a detailed analysis of their lightcurves and we were able to determine periods for 14 of these candidates. Of the 10 candidates without periods, we could rule out the transiting planet model for 7 of them by determining the minimum companion radius and for another one by predicting the orbital period. For INT-7789-TR-1, it was only at the  $1\sigma$  level that we could rule out the transiting planet model based on the shape of the best eclipse. For INT-7789-TR-2 we found a companion radius of  $0.185^{+0.009}_{-0.000} R_{\odot}$  ( $1.81^{+0.09}_{-0.00} R_J$ ) based on the analysis of the best eclipse. Follow-up observations (see below) will be required for both of these candidates in order confirm that INT-7789-TR-1 is an

eclipsing binary and in order to determine the nature of INT-7789-TR-2.

For the 14 transit candidates with well determined periods, we could rule out the transiting planet model for 4 of them from the detection of previously disguised secondary eclipses, and for 3 of them from the observation that the out-of-eclipse lightcurve data exhibit ellipsoidal variations and heating effects. One of the candidates is possibly a new cataclysmic variable with a long period (10.8 h) which could be a cluster member, worthy of follow-up observations in its own right. All of the 8 above mentioned candidates plus another 5 may be ruled out as having planetary companions by considering that the companion radius obtained from the full transit fit is greater than  $0.2R_{\odot}$ . For INT-7789-TR-3, none of the above arguments may be used to rule out the transiting planet model. However, on application of an eclipsing binary model to the lightcurve we find that the model consisting a pair of grazing K dwarf stars is  $\sim 7.8$  times more likely than the transiting planet model. This is by no means a definitive conclusion that INT-7789-TR-3 is an eclipsing binary since there is a non-negligible probability that the transiting planet model is still valid. Follow-up observations will be required to confirm that INT-7789-TR-3 is the type of eclipsing binary that we predict in this thesis.

Future photometric observations of the three transit candidates for which we could not rule out the transiting planet model with confidence should consist of time series observations in two different filters. Eclipsing binary status may be confirmed by the observation of different eclipse depths in different filters since a planetary transit is an achromatic event. INT-7789-TR-1 and INT-7789-TR-2 also require the observations of multiple eclipses in order to determine their period and whether they exhibit secondary eclipses or not. If these follow-up photometric observations still allow the possibility that the transiting planet model is valid, then radial velocity observations may be used to place an upper limit on the mass of the orbiting companion, hopefully low enough to rule out a stellar or brown dwarf companion. The fact that these candidates are so faint ( $r' \approx 20.7$  mag for INT-7789-TR-1,  $r' \approx 18.0$  mag for INT-7789-TR-2 and  $r' \approx 19.6$  mag for INT-7789-TR-3) makes it very unlikely that radial velocity observations with 10-m class telescopes will achieve the accuracy required to determine the actual mass of the companion (Charbonneau 2003).

The most important conclusion from this thesis is that our transit survey from the 2000-

09 run alone has just about reached the sensitivity required to detect a few hot Jupiters if the abundance of such planets in the field of NGC 7789 is similar to that of the Solar neighbourhood. At most we have detected 3 transiting hot Jupiters, but our analysis of these candidates shows that this is very unlikely. Follow-up observations will most likely show that our candidates are eclipsing binaries, which means that our transit survey will have produced a null result. Unfortunately it is not possible to estimate the hot Jupiter fraction for this field under the assumption that our survey has produced a null result. However we have been able to place useful limits on the hot Jupiter fraction of the different types of stars in our sample. The most stringent limit on the 3d to 5d hot Jupiter fraction was obtained for the G stars at 5.2% with a significance level of 1%. This was assuming a typical hot Jupiter radius similar to that of HD 209458b at  $\sim 1.4R_J$  (Mazeh et al. 2000). Extension of our Monte Carlo simulations to the 1999-07 run, where we also looked for transits, will serve to increase the number of transiting planets that we expected to detect, and assuming that the 3 transit candidates are shown to be eclipsing binaries by follow-up observations, then we will be able to place even better limits on the planet fraction as a function of star type for this field.

# A

---

## The $M \sin i$ Ambiguity

### Theorem 1:

Consider an extra-solar planet P in a circular orbit around a star S such that the orbit lies in a plane  $\Pi$  with normal  $\underline{\hat{n}}$ . If the orientation of the orbital plane  $\Pi$  is random, then the probability that the value of  $\sin i$  is greater than some value  $t \in \mathbb{R}$  is given by:

$$P(\sin i > t) = \cos(\sin^{-1}(t))$$

### Proof:

Without loss of generality, let  $\underline{\hat{n}}$  be inclined at an angle  $0 \leq i \leq \pi/2$  to the line of sight  $\underline{\hat{i}}$  (see Figure A.1):

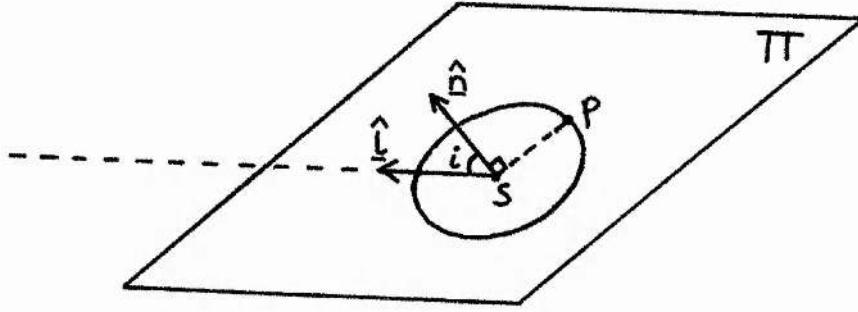


Figure A.1: General configuration of an extra-solar planet in a circular orbit.

The fact that the orientation of the orbital plane  $\Pi$  is random implies that the vector  $\hat{l}$  is distributed uniformly over the surface of a hemisphere [1].

The coordinates of the point A in the  $xy$ -plane as defined in Figure A.2 are given by:

$$x = r \cos \theta$$

$$y = r \sin \theta \quad [2]$$

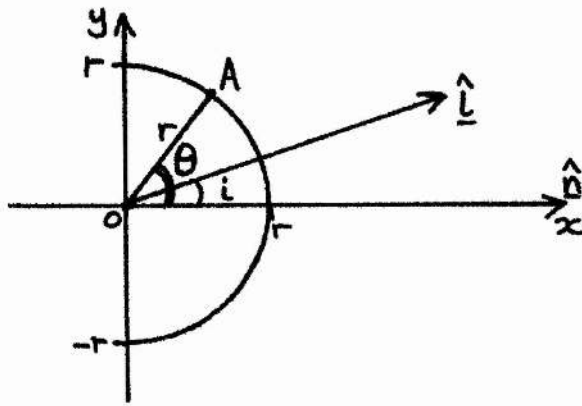


Figure A.2: Definition of the  $xy$ -plane for the extra-solar planet.



Let:

$S_1$  = Surface area of the hemisphere.

$S_2$  = Surface area of the cap on the hemisphere defined such that  $i \leq i_0$ .

Then:

$$P(i \leq i_0) = \frac{S_2}{S_1} \quad [3] \text{ (Using [1])}$$

Now:

$$S_1 = 2\pi r^2 \quad [4]$$

Also:

$$\begin{aligned} S_2 &= \int_0^{i_0} (2\pi y) r \, d\theta \\ &= 2\pi r^2 \int_0^{i_0} \sin \theta \, d\theta \quad \text{(Using [2])} \\ &= 2\pi r^2 \left[ -\cos \theta \right]_0^{i_0} \\ &= 2\pi r^2 (1 - \cos i_0) \quad [5] \end{aligned}$$

Substitute [4] & [5] into [3]:

$$P(i \leq i_0) = 1 - \cos i_0 \quad [6]$$

Now:

$$\begin{aligned} P(\sin i > t) &= 1 - P(\sin i \leq t) \\ &= 1 - P(i \leq \sin^{-1} t) \\ &= 1 - (1 - \cos(\sin^{-1} t)) \quad \text{(Using [6])} \\ &= \cos(\sin^{-1} t) \end{aligned}$$

**Example 1:**

The median value of  $\sin i$  for an extra-solar planet as defined in Theorem 1 is given by:

$$\begin{aligned} P(\sin i > t) &= 0.5 \\ \therefore \cos(\sin^{-1} t) &= 0.5 && (\text{By Theorem 1}) \\ \therefore t &= \sqrt{3}/2 \end{aligned}$$

Similarly, for the 95% and 99% lower limits to  $\sin i$  we have:

$$\begin{aligned} P(\sin i > t) &= 0.95 \\ \therefore t &= 0.31225 \dots \approx 0.312 && (\text{By Theorem 1}) \end{aligned}$$

and:

$$\begin{aligned} P(\sin i > t) &= 0.99 \\ \therefore t &= 0.14106 \dots \approx 0.141 && (\text{By Theorem 1}) \end{aligned}$$

### Example 2:

For an extra-solar planet of mass  $M_p$  with measured/projected mass  $M_p \sin i$ , calculate the value of  $\langle M_p \rangle$ .

Firstly:

$$\begin{aligned} P(0 \leq x \leq i_0) &= 1 - \cos i_0 && (\text{Using [6]}) \\ \therefore \frac{dP}{di_0} &= \sin i_0 && [7] \end{aligned}$$

Then:

$$\begin{aligned}\langle M_p \rangle &= \left\langle \frac{M_p \sin i}{\sin i} \right\rangle \\&= M_p \sin i \left\langle \frac{1}{\sin i} \right\rangle \\&= M_p \sin i \int_0^{\pi/2} \frac{1}{\sin i_0} \frac{dP}{di_0} di_0 \\&= M_p \sin i \int_0^{\pi/2} \frac{1}{\sin i_0} \sin i_0 di_0 \quad \text{(Using [7])} \\&= \frac{\pi}{2} M_p \sin i\end{aligned}$$

# B

---

## Useful Data/Constants

$$\odot \equiv \text{Sun}$$

$$\text{J} \equiv \text{Jupiter}$$

$$\oplus \equiv \text{Earth}$$

$$1 \text{ AU} \equiv 1.496 \times 10^{11} \text{ m}$$

$$1 \text{ pc} \equiv 3.086 \times 10^{16} \text{ m}$$

$$1 \text{ pc} \equiv 3.262 \text{ ly}$$

$$a_{\text{J}} = 5.203 \text{ AU}$$

$$a_{\oplus} = 1.000 \text{ AU}$$

$$R_{\odot} = 6.960 \times 10^8 \text{ m}$$

$$R_{\text{J}} = 7.131 \times 10^7 \text{ m}$$

$$R_{\oplus} = 6.378 \times 10^6 \text{ m}$$

$$M_{\odot} = 1.989 \times 10^{30} \text{ kg}$$

$$M_{\text{J}} = 1.899 \times 10^{27} \text{ kg}$$

$$M_{\oplus} = 5.974 \times 10^{24} \text{ kg}$$

$$R_{\text{J}} = 0.1025 R_{\odot}$$

$$M_{\text{J}} = 9.548 \times 10^{-4} M_{\odot}$$

$$G = 6.673 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$$

$$c = 2.998 \times 10^8 \text{ ms}^{-1}$$

---

# C

---

## Radial Velocity Curves

### Theorem 2:

Consider a star S orbiting the centre of mass C of a planetary system containing a single extra-solar planet. Let the orbit of the star lie in a plane  $\Pi'$  with normal  $\hat{n}$  and let the plane  $\Pi$  have a normal  $\hat{i}$  along the line of sight from the Earth. Without loss of generality, let  $\hat{n}$  be inclined at an angle  $0 \leq i \leq \pi/2$  to  $\hat{i}$ , and let the planes  $\Pi$  and  $\Pi'$  intersect along the line NN'. Also, without loss of generality, let us define a system of perpendicular axes  $xyz$  in  $\Pi$  and  $x'y'z'$  in  $\Pi'$  such that the origins O and O' are coincident at C, the  $x$  and  $x'$  axes both lie along NN', the  $z$  axis lies in the direction of  $\hat{i}$  and the  $z'$  axis lies in the direction of  $\hat{n}$ . Finally, let P and A denote the periastron and apastron respectively of the orbit of S around C, let  $\omega$  be the angle PCN, let  $\theta$  be the angle PCS and let  $r$  denote the distance from C to S. Graphically we have:



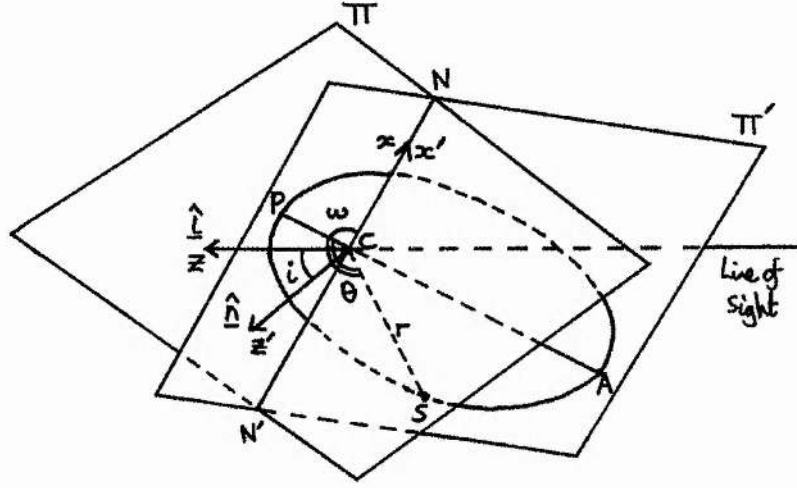


Figure C.1: General configuration of the orbit of a star about its centre of mass.

Then:

$$\frac{dz}{dt} = \left( \frac{2\pi G}{P} \right)^{1/3} \frac{M_p \sin i}{(M_* + M_p)^{2/3} (1 - e^2)^{1/2}} (e \cos \omega + \cos(\theta + \omega))$$

**Proof:**

The general position vector of S in  $\Pi'$  is given by:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} r \cos(\theta + \omega) \\ r \sin(\theta + \omega) \\ 0 \end{pmatrix} \quad [1]$$

The linear transformation from  $x'y'z'$  to  $xyz$  is a rotation through an angle  $-i$  around the

$x'$  axis. Such a linear transformation is represented by:

$$\begin{aligned} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(-i) & \sin(-i) \\ 0 & -\sin(-i) & \cos(-i) \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos i & -\sin i \\ 0 & \sin i & \cos i \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \end{aligned} \quad [2]$$

Substitute [1] into [2]:

$$\begin{aligned} z &= r \sin(\theta + \omega) \sin i \\ \therefore \frac{dz}{dt} &= \frac{dr}{dt} \sin(\theta + \omega) \sin i + r \frac{d\theta}{dt} \cos(\theta + \omega) \sin i \\ &= r \frac{d\theta}{dt} \sin i \left( \frac{dr}{d\theta} \frac{\sin(\theta + \omega)}{r} + \cos(\theta + \omega) \right) \end{aligned} \quad [3]$$

Let the orbit of S have semi-major axis  $a_*$  and eccentricity  $e$ . Then, for an ellipse:

$$\begin{aligned} r &= \frac{a_*(1 - e^2)}{1 + e \cos \theta} \\ \therefore \frac{dr}{d\theta} &= -a_*(1 - e^2)(1 + e \cos \theta)^{-2}(-e \sin \theta) \\ &= \frac{a_* e (1 - e^2) \sin \theta}{(1 + e \cos \theta)^2} \\ &= \frac{r e \sin \theta}{1 + e \cos \theta} \end{aligned} \quad [4] \quad [5] \text{ (Using [4])}$$

Substitute [5] into [3]:

$$\begin{aligned} \frac{dz}{dt} &= r \frac{d\theta}{dt} \sin i \left( \frac{e \sin \theta \sin(\theta + \omega)}{1 + e \cos \theta} + \cos(\theta + \omega) \right) \\ &= r \frac{d\theta}{dt} \frac{\sin i}{(1 + e \cos \theta)} (e \sin \theta \sin(\theta + \omega) + e \cos \theta \cos(\theta + \omega) + \cos(\theta + \omega)) \\ &= r \frac{d\theta}{dt} \frac{\sin i}{(1 + e \cos \theta)} (e \cos \omega + \cos(\theta + \omega)) \end{aligned} \quad [6]$$

Let  $M_*$  and  $M_p$  denote the masses of the star and the extra-solar planet respectively. Let  $P$  be the period of the star's orbit. Then Kepler's third law (Equation 1.8) states that:

$$P^2 = \frac{4\pi^2 a_*^3}{GM_c} \quad [7]$$

where  $M_c$  is the equivalent gravitational mass of the centre of mass C given by:

$$M_c = \frac{M_p^3}{(M_* + M_p)^2} \quad [8]$$

Substitute [8] into [7] and rearrange:

$$a_* = M_p G^{1/3} \left( \frac{P}{2\pi(M_* + M_p)} \right)^{2/3} \quad [9]$$

The total angular momentum of the star  $L_*$  is given by:

$$L_* = M_* \sqrt{GM_c a_* (1 - e^2)} \quad [10]$$

The law of conservation of angular momentum requires that:

$$\begin{aligned} M_* r^2 \frac{d\theta}{dt} &= L_* \\ \therefore r \frac{d\theta}{dt} &= \frac{L_*}{M_* r} \\ &= \frac{M_* \sqrt{GM_c a_* (1 - e^2)}}{M_* a_* (1 - e^2)} (1 + e \cos \theta) \quad (\text{Using [4] \& [10]}) \\ &= \sqrt{\frac{GM_c}{a_* (1 - e^2)}} (1 + e \cos \theta) \quad [11] \end{aligned}$$

Substitute [11] into [6]:

$$\begin{aligned} \frac{dz}{dt} &= \sqrt{\frac{GM_c}{a_* (1 - e^2)}} \sin i (e \cos \omega + \cos(\theta + \omega)) \\ &= \left[ \frac{(2\pi(M_* + M_p))^{1/3}}{M_p^{1/2} G^{1/6} P^{1/3}} \right] \left[ \frac{M_p^{3/2}}{(M_* + M_p)} \right] \frac{G^{1/2}}{(1 - e^2)^{1/2}} \sin i (e \cos \omega + \cos(\theta + \omega)) \\ &\quad (\text{Using [8] \& [9]}) \\ &= \left( \frac{2\pi G}{P} \right)^{1/3} \frac{M_p \sin i}{(M_* + M_p)^{2/3} (1 - e^2)^{1/2}} (e \cos \omega + \cos(\theta + \omega)) \end{aligned}$$

**Theorem 3:**

The quantity  $K$  defined by:

$$K = \left( \frac{2\pi G}{P} \right)^{1/3} \frac{M_p \sin i}{(M_p + M_*)^{2/3} (1 - e^2)^{1/2}}$$

is the semi-amplitude of the radial velocity curve for star S.

**Proof:**

The radial velocity  $dz/dt$  is at a maximum when  $\theta + \omega = 0$  and at a minimum when  $\theta + \omega = \pi$  (By Theorem 2). Denoting the maximum value of  $dz/dt$  by  $A$  and the minimum value of  $dz/dt$  by  $B$ , we have:

$$A = Ke \cos \omega + K \quad [12]$$

$$B = Ke \cos \omega - K \quad [13]$$

Subtract [13] from [12]:

$$\begin{aligned} A - B &= 2K \\ \therefore K &= \frac{A - B}{2} \end{aligned}$$

Hence  $K$  is the semi-amplitude of the radial velocity curve for star S.

For further information, see Aitken (1935) and Hilditch (2001).

## **Acknowledgements**

This research made use of the SIMBAD data base operated at CDS, Strasbourg, France and the WEBDA data base operated at the University Of Lausanne, Switzerland. This thesis was based on observations made with the Isaac Newton Telescope operated by the Isaac Newton Group on the island of La Palma in the Spanish Observatorio del Roque de los Muchachos. D.M.Bramich was funded by a PPARC research studentship during the course of this work, and would like to thank both St. Andrews University, Scotland, and the Instituto de Astrofisica de Canarias, Spain, for their hospitality during this time.

Personally I would like to thank my supervisor, Keith Horne, for the constant flow of enthusiasm and interesting ideas to work on. I will miss the 4pm tea breaks poring over fresh plots along with the many discussions from which I learned a great deal. A special kind of thanks goes to Rocio Muñiz Santacoloma who provided the inspiration and support necessary in order to complete this work. She was always there to put up with me throughout the whole three years, quite a feat of endurance. Finally Farton must be mentioned for his faithful company and self control in not eating my thesis...

# Bibliography

- Aitken, R. 1935, *The Binary Stars* (McGraw-Hill Book Company Inc.), 125
- Alard, C. 2000, *A&AS*, 144, 363
- Alard, C., & Lupton, R. 1998, *ApJ*, 503, 325
- Alcock, C., et al. 1993, *Nature*, 365, 621
- Alcock, C., et al. 1995, *ApJ*, 454, L125
- Baraffe, I., et al. 1998, *A&A*, 337, 403
- Baranne, A., Mayor, M., & Poncet, J. 1979, *Vistas Astron.*, 23, 279
- Baranne, A., et al. 1996, *A&AS*, 119, 373
- Bessell, M. 1990, *A&AS*, 83, 357
- Bond, I., et al. 2001, *MNRAS*, 327, 868
- Bond, I., et al. 2004, *ApJ*, 606, L155
- Boss, A. 1980, *ApJ*, 236, 619
- Boss, A. 1986, in *Astrophysics Of Brown Dwarfs*, 206
- Boss, A. 1987, *ApJ*, 319, 149
- Bouchy, F., et al. 2004, *astro-ph/0404264*
- Brown, T., et al. 1998, *ApJ*, 494, L85
- Brown, T., et al. 2001, *ApJ*, 552, 699
- Burrows, A., & Liebert, J. 1993, *Rev. Mod. Phys.*, 65, 301



- Butler, R., & Marcy, G. 1996, *ApJ*, 464, L153
- Butler, R., et al. 1997, *ApJ*, 474, L115
- Butler, R., et al. 1999, *ApJ*, 526, 916
- Butler, R., et al. 2000, in *Planetary Systems In The Universe: Observation, Formation And Evolution*, 1
- Cameron, A., et al. 2002, *MNRAS*, 330, 187
- Charbonneau, D. 2003, *astro-ph/0302216*
- Charbonneau, D., et al. 2000, *ApJ*, 529, L45
- Cochran, W., et al. 1997, *ApJ*, 483, 457
- Cody, A., & Sasselov, D. 2002, *ApJ*, 569, 451
- Drake, A. 2003, *ApJ*, 589, 1020
- Drake, A., & Cook, K. 2004, *ApJ*, 604, 379
- Draper, P. 2000, in *Astronomical Data Analysis Software And Systems IX*, 615
- Dreizler, S., et al. 2002, *A&A*, 391, L17
- Gaudi, B., et al. 2002, *ApJ*, 566, 463
- Gilliland, R., et al. 2000, *ApJ*, 545, L47
- Gim, M., et al. 1998, *PASP*, 110, 1318
- Goldreich, P., & Tremaine, S. 1980, *ApJ*, 241, 425
- Gonzalez, G. 1998, *A&A*, 334, 221
- Gray, D., & Hatzes, A. 1997, *ApJ*, 490, 412
- Hatzes, A., Cochran, W., & Bakker, E. 1998, *ApJ*, 508, 380
- Hatzes, A., Cochran, W., & Johns-Krull, C. 1997, *ApJ*, 478, 374
- Heacox, W. 1999, *ApJ*, 526, 928
- Hilditch, R. 2001, *An Introduction To Close Binary Stars* (Cambridge University Press)
- Ida, S., & Lin, D. 2004, *ApJ*, 604, 388

- Jones, H. 2002, in *Techniques For The Detection Of Planets And Life Beyond The Solar System*, 1
- Konacki, M., Maciejewski, A., & Wolszczan, A. 2000, *ApJ*, 544, 921
- Konacki, M., et al. 2003a, *ApJ*, 597, 1076
- Konacki, M., et al. 2003b, *Nature*, 421, 507
- Konacki, M., & Wolszczan, A. 2003, *ApJ*, 591, L147
- Lang, K. 1992, *Astrophysical Data: Planets and Stars* (Springer-Verlag)
- Leigh, C., et al. 2003a, *MNRAS*, 344, 1271
- Leigh, C., et al. 2003b, *MNRAS*, 346, L16
- Lin, D., Bodenheimer, P., & Richardson, D. 1996, *Nature*, 380, 606
- Lin, D., & Papaloizou, J. 1986, *ApJ*, 309, 846
- Lineweaver, C., & Grether, D. 2002, *Astrobiol.*, 2, 325
- Lineweaver, C., & Grether, D. 2003, *ApJ*, 598, 1350
- Lineweaver, C., Grether, D., & Hidas, M. 2003, in *Scientific Frontiers in Research on Extrasolar Planets*, 161
- Mallén-Ornelas, G., et al. 2003, *ApJ*, 582, 1123
- Marcy, G., & Butler, R. 1996, *ApJ*, 464, L147
- Marcy, G., et al. 1997, *ApJ*, 481, 926
- Mayor, M., & Queloz, D. 1995, *Nature*, 378, 355
- Mazeh, T., et al. 2000, *ApJ*, 532, L55
- Monet, D., et al. 2003, *AJ*, 125, 984
- Nakano, T. 1987, *MNRAS*, 224, 107
- Noyes, R., et al. 1984, *ApJ*, 279, 763
- Pepe, F., et al. 2004, *A&A*, 423, 385
- Perryman, M. 2000, *Rep. Prog. Phys.*, 63, 1209

- Robin, A., Cr    , M., & Mohan, V. 1992, *ApJ*, 400, L25
- Robin, A., et al. 2003, *A&A*, 409, 523
- Saar, S., Butler, R., & Marcy, G. 1998, *ApJ*, 498, L153
- Saar, S., & Fischer, D. 2000, *ApJ*, 534, L105
- Sadakane, K., et al. 2002, *PASJ*, 54, 911
- Santos, N., Israelian, G., & Mayor, M. 2004, *A&A*, 415, 1153
- Santos, N., et al. 2003, *A&A*, 398, 363
- Seaton, M. 1979, *MNRAS*, 187, 73
- Shu, F., Adams, F., & Lizano, S. 1987, *Ann. Rev. Astron. Astrophys.*, 25, 23
- Snodgrass, C., Horne, K., & Tsapras, Y. 2004, *MNRAS*, 351, 967
- Soderblom, D. 1983, *ApJS*, 53, 1
- Stetson, P. 1987, *PASP*, 99, 191
- Street, R., et al. 2002, *MNRAS*, 330, 737
- Street, R., et al. 2003, *MNRAS*, 340, 1287
- Stuart, A., & Ord, J. 1987, *Kendall's Advanced Theory Of Statistics*, Vol. I, 5th Edition (New York: Oxford Univ. Press)
- Tarter, J. 1986, in *Astrophysics Of Brown Dwarfs*, 21
- Tinney, C., et al. 2001, *ApJ*, 551, 507
- Torres, G., et al. 2004, *ApJ*, 609, 1071
- Trilling, D., et al. 1998, *ApJ*, 500, 428
- Udalski, A., et al. 1993, *Acta Astron.*, 43, 289
- Udalski, A., et al. 1994, *ApJ*, 436, L103
- Udalski, A., et al. 2002a, *Acta Astron.*, 52, 1
- Udalski, A., et al. 2002b, *Acta Astron.*, 52, 115
- Udalski, A., et al. 2003, *Acta Astron.*, 53, 133

- Udry, S., et al. 2000, *A&A*, 356, 599
- Udry, S., Mayor, M., & Santos, N. 2003, *A&A*, 407, 369
- Wallace, P. 1998, Starlink User Note 5.17, Rutherford Appleton Laboratory
- Walton, N., et al. 2001, *ING Newsletter*, 4, 7
- Wambsganss, J. 1997, *MNRAS*, 284, 172
- Ward, W. 1997, *ApJ*, 482, L211
- Wolszczan, A. 1994, *Science*, 264, 538
- Wolszczan, A., & Frail, D. 1992, *Nature*, 355, 145
- Zeilik, M., & Gregory, S. 1998, *Introductory Astronomy & Astrophysics 4th Edition*  
(Saunders College Publishing)
- Zucker, S., & Mazeh, T. 2002, *ApJ*, 568, L113